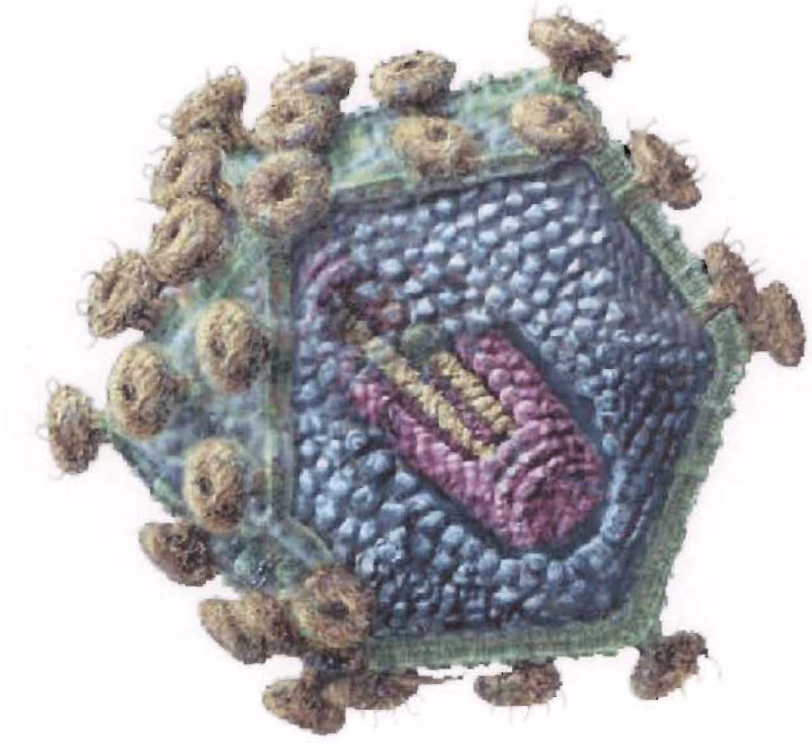
Studies on Natural Product Derivatives:

HIV Therapies Incorporating Marine Natural Products



A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Chemistry at the University of Canterbury by

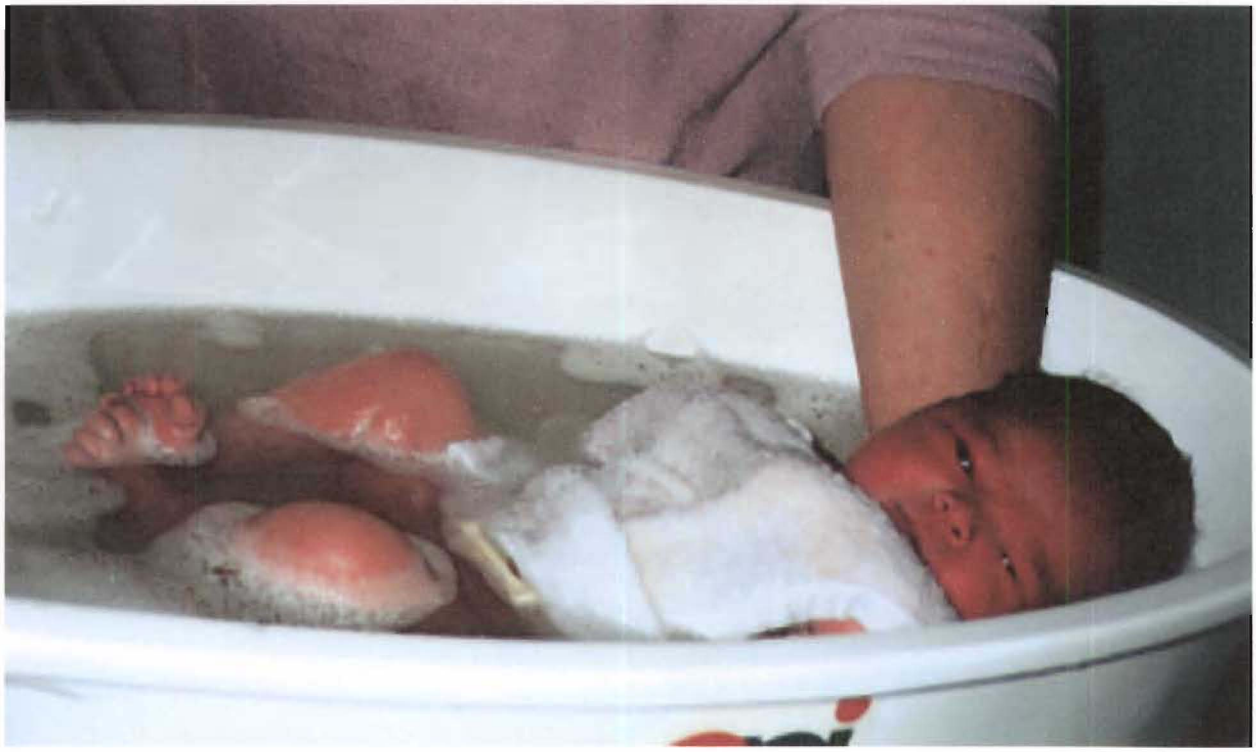
Scott D. Bringans

University of Canterbury, July, 2001

This thesis is dedicated to

Jack Alexander Bringans

born on the 5th of July, 2001.



Acknowledgments

I would like to express my thanks to my supervisors, Professor John Blunt and Professor Murray Munro, for providing such an interesting and stimulating project and for their encouragement and support both academically and personally. They have provided an excellent environment for learning for which I am grateful.

The technical staff at the University of Canterbury have provided excellent support and my thanks are expressed for both their freindship and assistance. Particular mention must be made to Bruce Clark for the countless hours spent teaching, helping and fixing problems with mass spectrometric analyses thoughout my time here. The skills I have attained are of invaluable importance and I am grateful for the input to my project. I would also like to thank Dr Lewis Pannell (NIDDK, NIH) for additional mass spectrometric analysis.

I am particularly grateful to Dr Michael Boyd, for firstly the collaborative agreement to undertake this project, and secondly for the invitation and financial assistance for my two month visit to the LDDRD, NCI. I would also like to thank Dr Toshi Mori for his supervision and kindness to me while at the LDDRD and to Dr Kirk Gustafson, Dr Barry O'Keefe and Dr Jim McMahon for their input, assistance and friendliness throughout my visit. Further thanks go to the numerous others working at the LDDRD who helped me and made me feel at home, particularly Shannon Berg without whose patience I would not have got very far at all.

Thanks must also go to the Canterbury Branch of the Royal Society and to the NZIC for contributions towards attendance at the New Zealand Institute of Chemistry Conference, Wellington 1999, and to the 4th International Symposium on Polymer Therapeutics, London 2000.

My fellow labmates have been a constant source of amusement, friendship, hilarity, intense juvenile behaviour and all round fun. Mention should be made of David Stirling, Rachel Lill, Sarah Hickford and Gill Nicholas for 'the formative years' and Sean Devenish, Warren MacLean, Marie Squire, Sylvia Urban, Greg Johnson, Jasmine Mahalm, Tim Harwood, Jonathan Hill, Hsini Chang, Richard Phipps, Alison Daines, Carol Dean, James Gardiner and all the others who have 'contributed' in their own special way over the last few years. Thanks for 'keeping it real'.

To my parents and family who never really understood what I do but think its sounds pretty cool anyway and have supported me in my endeavours, I thank you.

Finally to my ever-patient wife Sandra who has had to put up with my years of relentless academic study, my absences when working late and my promises that I'll be finished 'soon'. Thank you for your encouragement, support, understanding and love. Most of all thanks for carrying our baby for nine months plus!. I couldn't have done it without you (the baby and the Ph.D.!!)

Abstract

CV-N is an 11 kDa, anti-HIV protein that binds strongly to the envelope glycoprotein, gp120, expressed on the outer surface of the free virion and also on HIV-infected cells. As such, it represents an important lead for development of anti-HIV therapeutics. Marine toxins such as the halichondrins have potent *in vivo* cytotoxicities and are lethal to cells. The combination of this potency of the marine toxins with the unique targeting capability of CV-N has been harnessed to produce conjugates that have the potential to selectively target and eliminate HIV-infected cells.

Three forms of the protein were developed; the native protein itself, a derivative recombinantly produced in *E. coli* with an extra cysteine at the *C*-terminal (CV-N-Cys) and CV-N with the lysine side chain amines converted into thiols (thiolated-CV-N).

To facilitate release of the toxin within infected cells an enzymatically-cleavable pH-dependent biolinker was incorporated separating the toxin from the protein. The chemistry required for incorporation of protein, biolinker, and toxin, was established through synthesis of fluorescently labelled conjugates capable of reaction with CV-N. Biological testing of these derivatives showed no interference with the anti-HIV activity of the CV-N when conjugated in these model compounds. Synthetic strategies were developed to produce two derivatives of norhomohalichondrin B amine, both containing the cleavable biolinker, but with activation from succinimidyl esters and maleimido groups respectively.

Native CV-N was reacted with the succinimidyl ester derived toxin construct to produce a CV-N-biolinker-toxin conjugate. The maleimido derivative toxin construct was reacted with both CV-N-Cys and thiolated-CV-N to produce closely related CV-N-toxin conjugates. Investigations into the binding properties and cell toxicities of these conjugates is currently underway.

Table of Contents

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

Chapter 8

Abbreviations

Ab	antibodies
AIDS	Acquired Immune Deficiency Syndrome
BODIPY	4,4-difluoro-4-bora-3a-diaza-s-indacene
BSA	bovine serum albumin
COSY	correlation spectroscopy
CV-N	cyanovirin-N
d	doublet (NMR)
Da	daltons (in mass spectrometry)
DCC	dicyclohexylcarbodiimide
DCM	dichloromethane
^0C	degrees Celsius
δ	chemical shift in parts per million (NMR)
DIPEA	diisopropylethylamine
DMAP	4-(dimethylamino)pyridine
DMF	dimethylformamide
DSC	disuccinimidyl carbonate
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
ELISA	enzyme linked immunosorbent assay
equiv.	equivalent(s) (with respect to moles)

ESI	electrospray ionisation (in mass spectrometry)
FAB	fast atom bombardment (in mass spectrometry)
Fmoc	9-fluorenylmethoxycarbonyl
g	gram(s) or acceleration of gravity
HBTU	*O*-benzotriazol-1-yl-tetramethyluronium Hexafluorophosphate
HIV	Human Immunodeficiency Virus
HMBC	heteronuclear multiple bond coherence (NMR)
HOBt	1-hydroxybenzotriazole
HPLC	high-performance liquid chromatography
HR	high-resolution (in mass spectrometry)
HSQC	heteronuclear single quantum coherence (NMR)
Hz	Hertz
IPA	isopropyl alcohol
IPTG	isopropyl β-D-thiogalactopyranoside
J	coupling constant (in NMR)
L	litre(s)
LCMS	liquid chromatograph mass spectrometry
LDDRD	Laboratory for Drug Discovery, Research and Development
m	multiplet (NMR) or milli
M	moles per litre
μ	micro
m/z	mass to charge ratio (mass spectrometry)

mm	millimetre(s)
MS	mass spectrometry
MW	molecular weight
NHS	*N*-hydroxysuccinimide
nm	nanometre(s)
NMR	nuclear magnetic resonance
OD	optical density
PBS	phosphate buffered saline
PDA	photodiode array (HPLC)
ppm	parts per million (NMR)
s	singlet (NMR)
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SPPS	solid phase peptide synthesis
t	triplet (NMR)
TES	triethylsilane
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TLC	thin layer chromatography
UV	ultraviolet

Chapter 1

INTRODUCTION

1.1 HIV

1.1.1 HIV, AIDS – A Global Perspective

As of December 2000 the World Health Organisation (WHO) put the number of people infected with the HIV virus at 36.1 million. This is more than 50% higher than the 1991 projection using the data available at the time. The following snapshots of the effect of the virus demonstrate the impact this epidemic has had around the world.[1]

- Since the beginning of the epidemic 21.8 million people have already died, 4.5 million of them children under 15 years of age.
- In the year 2000, 5.3 million people became newly infected while 3 million died.
- In all parts of the world except sub-Saharan Africa, there are more men infected with HIV than women.
- During 2000, more new infections have been registered in the Russian Federation than in all previous years of the epidemic combined. The total number of people living with HIV/AIDS in Eastern Europe and countries of the former Soviet Union was 420,000 in 1999 but after 2000 is estimated to be 700,000. Intravenous drug use is the prevalent method of transmission in this region.
- For the first time the number of new infections in sub-Saharan Africa has declined with 3.8 million infected in 2000 compared to 4 million the year before. However, this region contributes 25.3 million of the global 36.1 million people infected. The prevalence rate for adults 15-49 years old living with HIV/AIDS in this area is 8.8%.

- In South Africa this epidemic is projected to reduce economic growth by 0.3-0.4% annually and by 2010 the GDP will be 17% lower than it should have been, wiping US$ 22 billion off the country's economy.
- In Abidjan AIDS costs represented between 0.8% and 3% of the wage bill in 1997. A survey of five Ethiopian firms in the mid 1990's found AIDS responsible for more than half the burden of sickness over a five-year period. In Tanzania a similar survey showed a three-fold increase in medical costs per employee while burial costs for the companies increased five-fold.
- While the prevalence rate for East Asia and the Pacific is only 0.07% there is ample room for growth. The use of illicit drugs and the sex trade are extensive, as is the migration and mobility within and across borders. China, which virtually wiped out sexually transmitted diseases in the 1960's, is now seeing a steep rise that could translate into higher HIV levels in future years. The hundred million or more people on the move within the country may exacerbate this situation.
- The Caribbean has the highest rate of infection outside of Africa with a 2.3% incidence.
- For Western Europe and North America the new infection rate has stalled while the prevalence rate has risen slightly in both regions mainly due to anti-retroviral therapies that are keeping HIV infected people alive longer. In an era where few young gay men have seen friends die of AIDS, and some mistakenly view anti-retrovirals as a cure, there is growing complacency about the HIV risk.

An analysis of the above facts reveals that the devastation wrought by HIV is very real, whether measured by the future prospects of a child or the bottom line of a company's ledger. Hundreds of millions of research dollars around the world continue to be spent on developing therapies to control the virus and on vaccines to prevent infection. The aim of this thesis is to contribute to this global effort so the 21st century may provide solutions to this devastating epidemic that has plagued the latter half of the 20th century.

1.1.2 The HIV Virus

The human immunodeficiency virus (HIV) is a member of the lentivirus genus of the *Retroviridae* family. A major feature characterising the *Retroviridae* family, is that rather than house DNA as their store of genetic material, retroviruses utilise RNA as their source of genetic information. Lentiviruses have certain characteristics that distinguish them from other retroviruses. The most prominent morphological difference is the cylindrical or cone shaped nucleoid in the mature virion. Other differences include the encoding of several extra genes not present in other retroviral genomes and a biphasic course of viral gene expression. One of the first viruses found in nature was a lentivirus, the equine infectious anaemia virus, which was discovered in 1904.[2] Lentiviruses infect a diverse group of animals including horses, sheep, goats, cows, cats, and primates.[3, 4] Within this group of animals the lentiviruses cause a wide variety of diseases. HIV-1 and HIV-2 are the only known human lentiviruses. Of the two subtypes, HIV-1 is the predominant form of the virus found throughout the world, while HIV-2 has been primarily isolated from West Africa, the Caribbean and South America. Both agents are associated with the development of progressive immunologic deterioration, but HIV-2 tends to have a longer incubation period[5] and is not as easily transmitted perinatally as HIV-1.[6]

The earliest documented case of AIDS known to have been caused by the HIV virus was in 1959 in a man from Kinshasa, Democratic Republic of Congo. However, it was not till the early 1980's that doctors in Los Angeles and New York were beginning to notice a pattern of rare pneumonia, cancer and other unfamiliar illnesses amongst young gay male patients.[7-9] It was noted that all of the illnesses documented were rare in people who have healthy immune systems. The search for the cause of AIDS focused on a wide variety of viruses including retroviruses but also herpesviruses and parvoviruses that were known to cause immune deficiency.[10, 11] Even after the discovery of HIV in 1983 it still took a year for its classification as a lentivirus.[12-15]

1.1.3 HIV Virion Structure

Each HIV virion is approximately 100 nm in diameter. Externally the surface usually consists of 72 knobs containing trimers of the envelope glycoproteins.[16] There are two extensively investigated surface glycoproteins, gp120 and gp41. These are derived from a single precursor of MW 160,000 that is cleaved inside the infected cell into the gp120 external surface envelope protein and the gp41 transmembrane protein. As mentioned, on the surface of the virion these are present in a trimeric form of gp120 with a single gp41

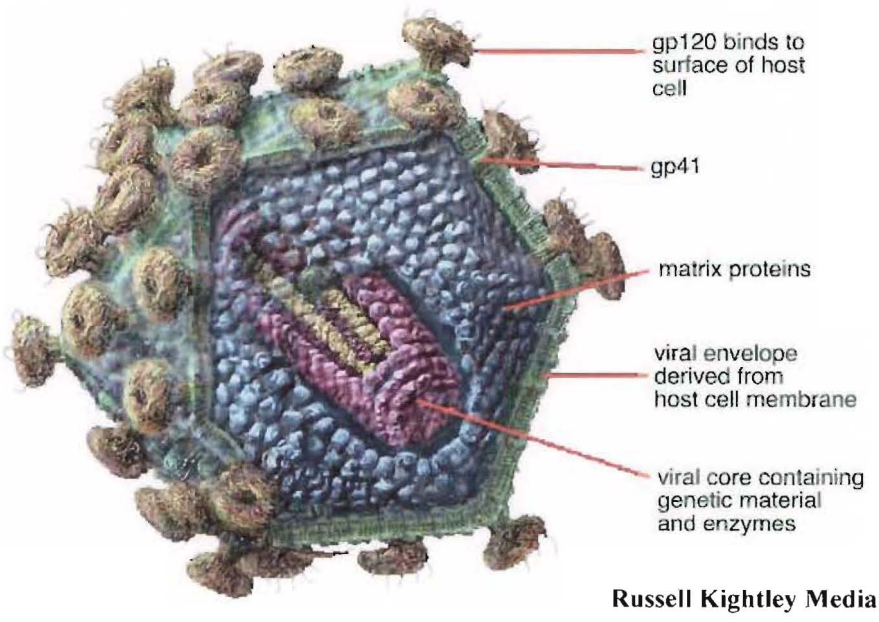


Figure 1.1 Stylised HIV virion structure

molecule buried between the three gp120 proteins. The gp120 protein is comprised of five variable regions (V1-V5) that are exposed on its surface and are interspersed with five conserved regions. Gp120 remains attached to gp41 through a series of non-covalent interactions located in both the *N*- and *C*-termini. The inner portion of the viral membrane is surrounded by a myristoylated p17 core (Gag) protein that provides the matrix for the viral structure.[17, 18] Within this is the p25 (or p24) capsid (Gag) protein that contains two identical RNA strands with which the DNA polymerase (p64/p51, also called Pol or reverse transcriptase) and nucleocapsid proteins (p9, p6) are closely associated. Overall the

virion has approximately 100 times more p25 than gp120, and 10 times more p25 than polymerase.[19]

1.1.4 Origins of HIV

It is generally agreed by most (but not all) that HIV-1 and HIV-2 represent novel, zoonotic introductions to the human populations within the last 100 years.[20] The most closely related animal lentivirus to HIV-1 is SIVcpz of chimpanzees. Of the SIVcpz genomes sequenced, the subspecies *Pan troglodytes troglodytes* is the closest to HIV-1 in genetic makeup.[21] Within the human population there are three major groups of HIV-1, coded M, N, and O. The group M represents all the clades (subtypes) from A to H that have spread to cause the worldwide pandemic.[22] The related subgroups O and N are by contrast largely confined to Cameroon, Gabon and neighbouring countries which happen to be the same areas as the natural habitat of *P. t. troglodytes*.[23] A comparison of the gene sequences of groups M, N, and O, show that they are as distinct from each other as they are from SIVcpz. (**Figure 1.2**)[21] This would tend to suggest that they have derived from three

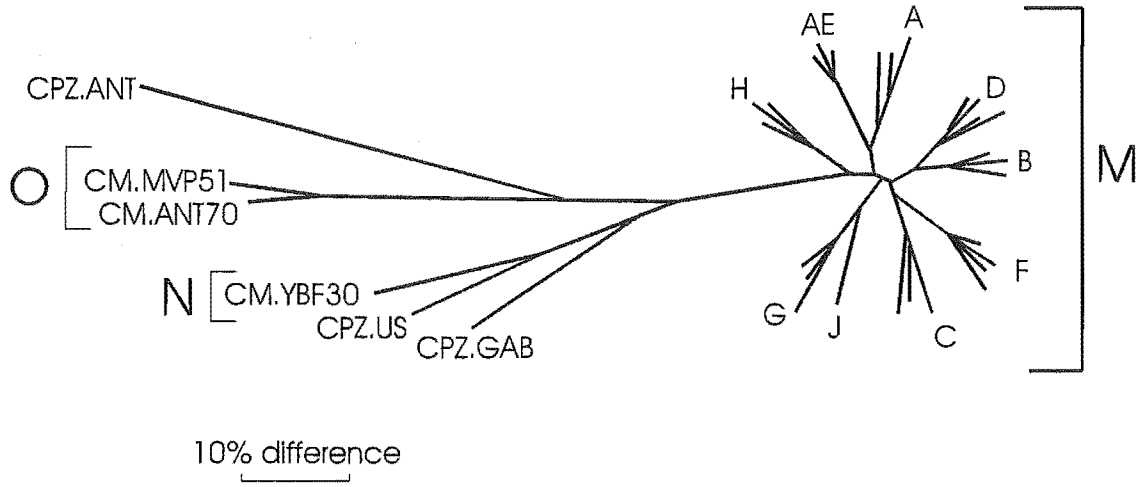


Figure 1.2 Genetic variability of HIV-1 subtypes

separate introductions from chimpanzees to humans, but only one has become pandemic.[24] The less common HIV-2 can also be subdivided into a number of major groups which appear to represent separate zoonoses from a different primate host, the West African sooty

mangabey which is infected with SIVsmm. There are many African monkeys that can harbour SIV strains but so far little evidence that they have crossed to humans. For example, despite 35 years use of unscreened green monkey kidneys to propagate poliovirus for the live attenuated vaccine, SIVagm has not crossed to humans. There is a hypothesis that HIV emerged from contaminated polio vaccine but the link appears very tenuous.[25]

1.1.5 Pathogenesis of HIV

The major reason why HIV is so problematic is that the target is the CD4 antigen which is present on several important cell types, particularly the CD4+ lymphocytes (also known as T-helper cells or T4 or CD4 T-cells) that are crucial to the body for a variety of host-immune responses. Over time the virus leaves the immune system with a number of CD4+ lymphocyte cells that do not function effectively. Depletion of these cells leads to a progressive immunoincompetence which has the dual effect of not only impairing the host's ability to inhibit HIV replication but also allows the emergence of opportunistic infections which characterise the onset of AIDS. An average person will have a T4 cell count of between 400 and 1200 in 1 mL of blood. Once the count falls below 200, the risk of opportunistic infections greatly increases. Without treatment, the average time from becoming HIV-infected to developing immune deficiency (AIDS) is around 10 years.

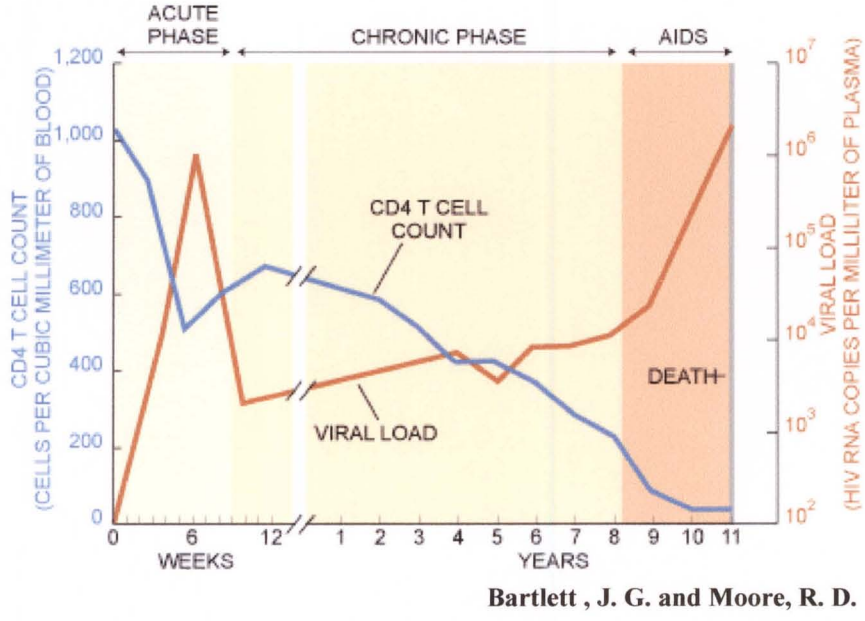


Figure 1.3 CD4 T cell count *vs* HIV viral load

As well as measuring the levels of CD4+ lymphocytes in the blood there is an assay that measures the amount of HIV-1 RNA present in the plasma. Currently, there are three commercially available varieties of this assay technique. These are the reverse transcriptase-polymerase chain reaction (RT-PCR), the branched DNA (bDNA), and the nucleic acid sequence-based amplification (NASBA).[26] The relationship between average viral load and the time from seroconversion to death is remarkably constant.[27] A study of infected individuals with identical CD4+ lymphocyte cell counts found that those with a high HIV-1 RNA plasma level died more rapidly (mean of 6.8 years) than those with a low level (time to death >10 years).[28] However, it has been estimated that only 2% of HIV is circulating in the blood.

During all phases of infection the average daily production of new HIV-1 virions is on the order of 1×10^{10} particles.[29] This high level of output of virions contrasts with the estimated $1\text{-}2 \times 10^{9}$ CD4+ lymphocyte cells eliminated per day.[30, 31] This in effect translates to 1% of the body's complement of CD4+ lymphocyte cells being destroyed and replaced every day over a period of many years. Another cell type that expresses CD4 antigens is the tissue macrophage, derived from monocytes circulating in the blood. These macrophages act as antigen-presenting cells and function as scavengers, but once infected act as an important reservoir of virus in the body.[32] It is the infection of these macrophage(s) that most probably accounts for the wasting syndrome in AIDS due to aberrant signalling of short range molecules such as cytokines and chemokines that traffic between different blood and tissue cells. Dendritic cells, which are derived from monocytes, also have CD4 antigens and so also become infected with HIV. After initial infection it is the dendritic cells that carry HIV to the lymph nodes where CD4+ lymphocytes become infected.[32] Using quantitative image analysis, a virus pool of 10-40 times larger than the pool of infected CD4+ lymphocytes has been estimated to be trapped on the surface of follicular dendritic cells (FDC).[33] This pool may be important for the spread and perpetuation of the viral infection within lymphoid tissue.

While the CD4 receptor is important for HIV attachment to the cell surface it was recognised early on that although it is necessary, it is not sufficient for cellular entry.[34]

After much research by many groups, there finally emerged not one but two co-receptors that mediate fusion between virion and cell. They are CXCR4[35] (also known as fusin) and CCR5 (also known as CC-KR-5).[36, 37] The CXCR4 receptor is utilised by syncytium-inducing (SI) primary HIV-1 isolates and laboratory-adapted HIV-1 strains. This variety of virus is termed T-tropic or X4,[38] and replicates well in primary T-cells and immortalised T-cell lines. In contrast, most non-syncytium-inducing (NSI) HIV-1 primary isolates use the CCR5 chemokine receptor and replicate within both macrophages and primary T cells but not in immortalised cell lines. For this reason these are referred to as macrophage (M)-tropic or R5 viruses. Most HIV strains of each subtype use CCR5 as their co-receptor but in about 50% of people that progress to AIDS the more virulent X4 strain emerges later in the course of the disease. This extra virulence is what probably hastens the depletion of CD4+ cells and emergence of AIDS symptoms. There are certain people that carry defective CCR5 receptors[39, 40] and about 1 in 400 of the Caucasian population (but rare in Africans) are homozygous for the 32 base-pair deletion.[41] Such people are highly resistant to infection and those who do become infected, do so through X4 strains that utilise the CXCR4 receptor.[42]

1.1.6 Life Cycle of HIV

The life cycle of HIV (**Figure 1.4**) begins at the surface of the host cell when the gp120 glycoprotein coating the virus binds to the CD4 receptor that is present on many lymphocytes. This binding is also facilitated by non-specific interactions between envelope proteins and other cell surface molecules such as heparan, sulfate proteoglycan, galactosyl ceramide, mannose receptors and adhesion molecules.[43] A very detailed understanding of just how gp120 interacts with CD4 was afforded with the crystal structure of a trimolecular complex consisting of the core of gp120, the V1/V2 domains of CD4 and the Fab fragment of a neutralising monoclonal antibody (mAb) designated 17b.[44] An analysis of this crystal structure reveals that the core of gp120 is composed of an inner and outer domain joined by a β-sheet, referred to as the bridging sheet.[44] The inner domain faces the axis of the trimer and presumably gp41, while the outer domain is exposed on the

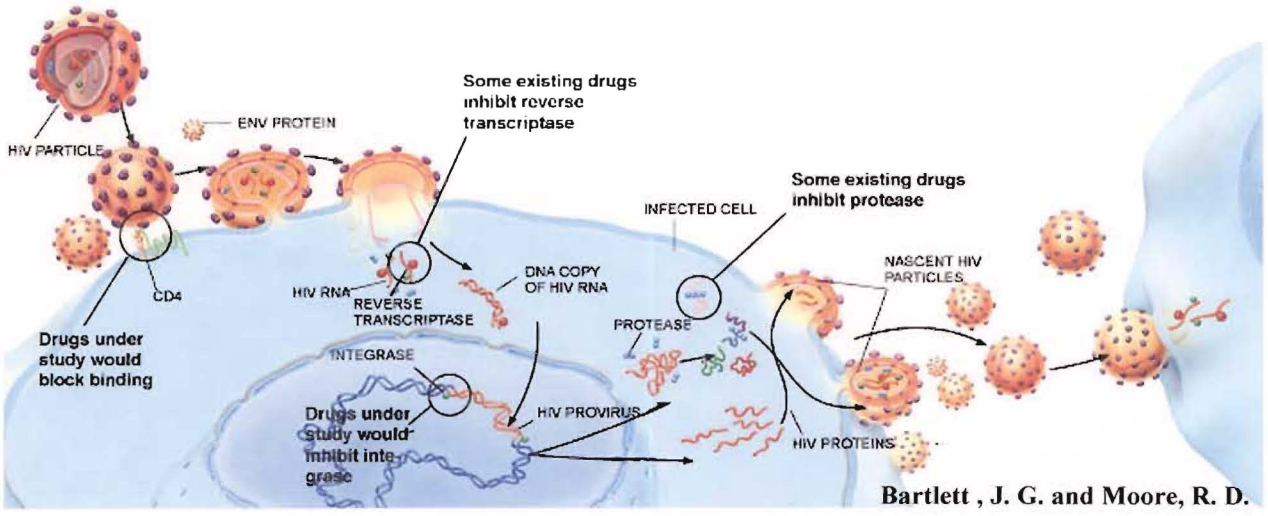


Figure 1.4 HIV life cycle

surface.[44] The CD4 receptor binds to gp120 in a recessed cavity constructed from elements of both major domains and the bridging sheet.[44] This deep pocket, which is occupied by the aromatic side chain of Phe43 of CD4, is very important for gp120 binding.[44] Unlike other portions of gp120 this cavity is highly conserved making inhibitors of this pocket a highly attractive target. Previous, less specific gp120 inhibitors have shown poor potency against diverse clinical isolates, probably due to the vast genetic diversity present in the HIV envelope protein.

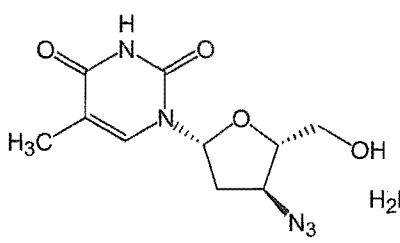
The binding to CD4 induces a conformational change within gp120, allowing interaction with either of the two chemokine receptors (CCR5 or CXCR4) to form a trimolecular complex (CD4-gp120-CCR5/CXCR4). An excess of the ligands (natural or synthetic) for these receptors can competitively inhibit this step.[45-48] AMD-3100, a bicyclam analogue, that inhibits the CXCR4 chemokine receptor is one of these promising ligands and has entered phase II trials in 2000 and will be launched in 2003 as a potential treatment for HIV infection if trials are successful.[49]

After binding to the chemokine receptors, the gp120 is thought to be stripped off the virion, exposing a hydrophobic domain at the *N*-terminus of gp41 which in turn mediates fusion of the virus and host cell membranes, therefore allowing the viral core access to the host cell cytoplasm. The genetic material of the HIV, which like all retroviruses is RNA, is now

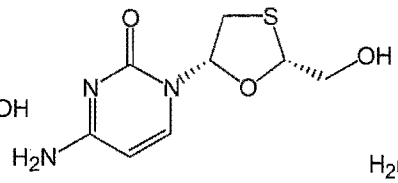
released and undergoes reverse transcription into double stranded DNA. The enzyme that performs this transformation (reverse transcriptase) is the target of several anti-retroviral drugs that are now in clinical use. These reverse transcriptase inhibitors will be discussed later. Now that the genetic material has been converted into DNA it enters the nucleus of the host cell where it is spliced (integrated) into the genetic material of the cell by the enzyme integrase. Inhibitors of this enzyme are also being developed as anti-retroviral therapies.[50] At this stage it is possible that the HIV may persist in a latent state for many years. These latently infected cells are the major barrier to eradication or cure of HIV. Activation of these cells using the integrated DNA, or provirus, leads to the transcription of viral DNA into messenger RNA (mRNA), which is then translated into viral proteins and RNA. The viral RNA and viral proteins now gather at the cell membrane and begin to form a new virus. Compounds are being developed that target the highly conserved HIV-1 nucleocapsid protein zinc fingers involved in genome packaging and virus assembly.[51, 52] A third enzyme, HIV protease, cleaves some of the viral proteins into their functional forms. Protease inhibitors block this crucial maturation step of the life cycle. Following assembly at the cell surface, the new viral particle buds from the cell, ready to infect others.

1.1.7 Current and Future Therapies

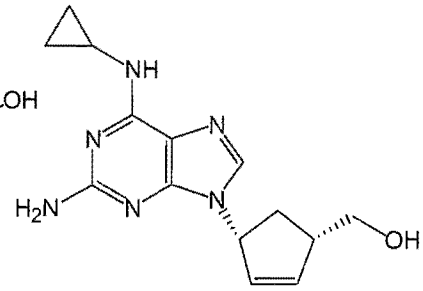
There are currently fourteen compounds that have been approved for the treatment of HIV infections. Of the fourteen, six are nucleoside reverse transcriptase inhibitors (NRTI's) that are competitive inhibitors of the normal substrates (dNTP's) after their intracellular conversion to the 5'-triphosphate form. These are zidovudine (AZT) (**1**), lamivudine (3TC) (**2**), abacavir (ABC) (**3**), zalcitabine (ddC) (**4**), didanosine (ddI) (**5**), and stavudine (d4T) (**6**). There are three non-nucleoside reverse transcriptase inhibitors (NNRTI's) that interfere with reverse transcriptase at a nonsubstrate binding, allosteric site. They are nevirapine (**7**), delavirdine (**8**) and efavirenz (**9**). The final class are the protease inhibitors (PI's) that prevent the cleavage of precursor to mature HIV proteins and therefore weaken the infectivity of the virions produced when these inhibitors are present. The aspartyl protease responsible for the proteolytic cleavage has been examined at the atomic level[53]



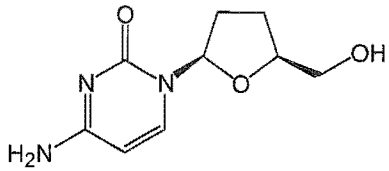
zidovudine (**1**)



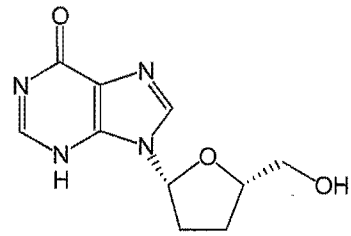
lamivudine (**2**)



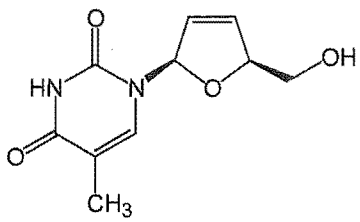
abacavir (**3**)



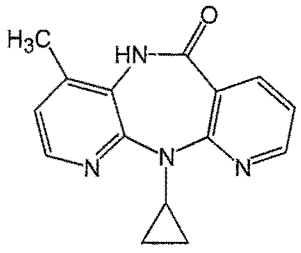
zalcitabine (**4**)



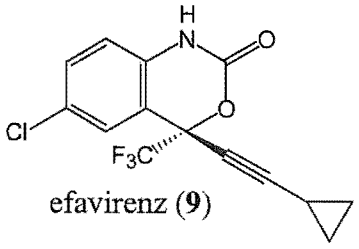
didanosine (**5**)



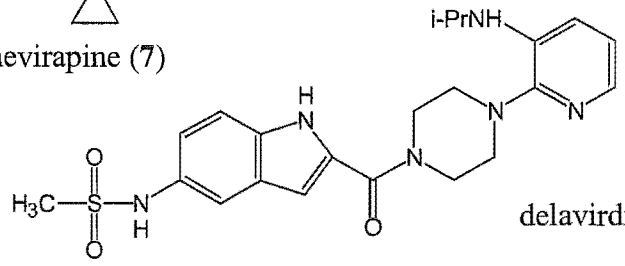
stavudine (**6**)



nevirapine (**7**)

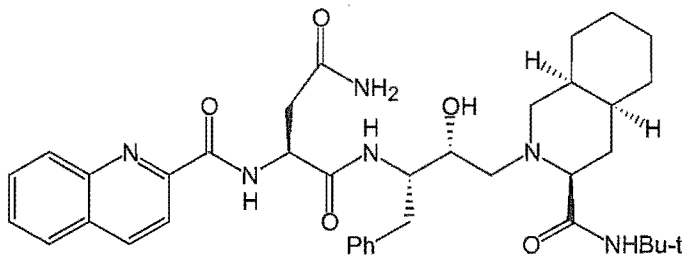


efavirenz (**9**)

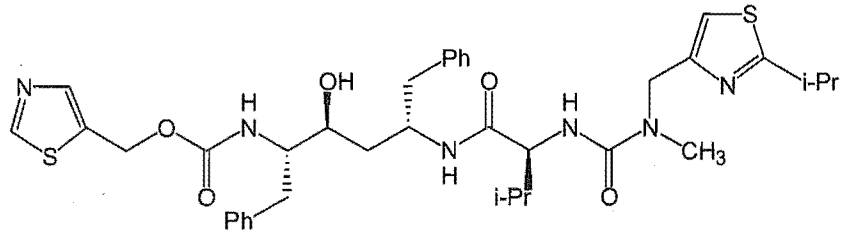


delavirdine (**8**)

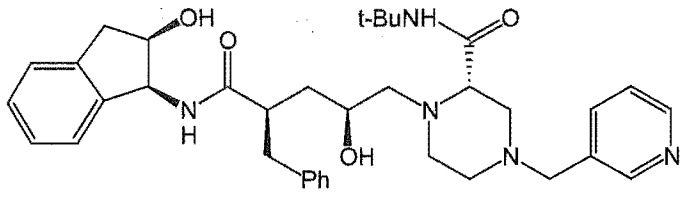
and subsequent rational design of inhibitors took place.[54] The five PI's are saquinavir (**10**), ritonavir (**11**), indinavir (**12**), nelfinavir (**13**) and amprenavir (**14**). All of the fourteen compounds are synthetic and not derived from natural sources.



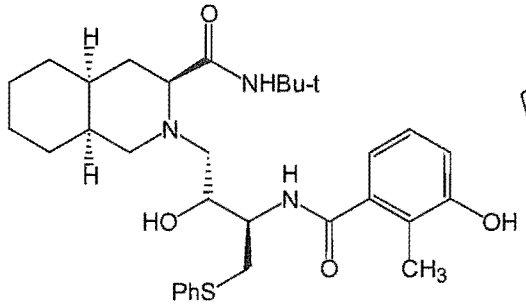
saquinavir (**10**)



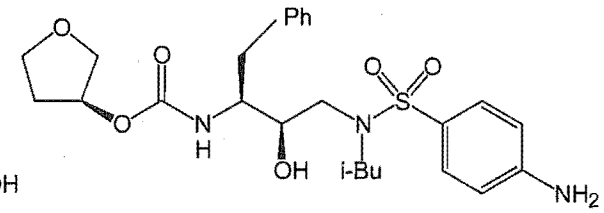
ritonavir (**11**)



indinavir (**12**)



nelfinavir (**13**)



amprenavir (**14**)

Drug resistance became a problem when some of the earlier drugs, above, were taken singly. As HIV replicates itself, mutations occur in a small percentage of new virus particles formed. Out of the mutated viruses there may be some that are unaffected by the particular anti-retroviral in use. The original viruses will die off but the new drug resistant mutated form will continue to replicate and the viral load will again increase. Therefore a new strategy, HAART (Highly Active Anti-Retroviral Therapy), was developed to try and overcome the growing drug resistance.[55] Studies clearly showed that a combination of

three drugs ('triple cocktail') were much more effective than one drug used alone or two drug combinations.[56, 57] This HAART strategy is now the standard treatment for HIV infection in the western world. It was thought that this therapy might have lead to a 'cure' if HIV could be completely eliminated but unfortunately viremia can be reactivated following cessation of treatment.[58] There is also the problem of imperfect adherence to the prescribed regimen of complex timetables and the often severe side effects of current therapies. This can lead to the increasing resistance of the virus particularly if only one drug is taken while another is not, due to deleterious side effects.[59]

Another characteristic, crucial to survival of the HIV-1 virus, that has been hinted at earlier, is the large reservoir of latently infected cells present in the body. Virus production is a dynamic process involving continuous rounds of infection of, and replication in, CD4+ T cells with a very rapid turnover of both virions and infected cells. With effective anti-retroviral therapy this process is largely, but not completely interrupted and after several months plasma levels of HIV RNA become undetectable in many patients. However, there are several potential reservoirs, both cellular and anatomically that may contribute to long-term persistence of HIV-1.[60] These include areas of the central nervous system and the male urogenital tract. By far the most troublesome reservoir is that of the latently infected resting memory CD4+ T cells carrying integrated HIV-1 DNA.[60] With a half-life of 44 months, this reservoir would take over 60 years of continuous treatment with conventional anti-retrovirals to be eliminated.[58] With the side effects and toxicities of current therapeutics this is a completely unrealistic hope.

A potential strategy for eliminating the virus along with the latently infected cells is to force the reservoir into a fully productive state, and to therefore be able to eliminate it with current drugs. This has already been investigated with the immune-activating compound interleukin-2.[61, 62] An alternative approach is to use a cytoreductive treatment such as cyclophosphamide that may reduce the size of the latent HIV-1 reservoir and thereby facilitate its elimination.[63]

1.1.8 HIV Vaccine Development

Since 1987, more than 40 individual vaccines for HIV have been studied in clinical trials worldwide. The US has committed itself to developing such a vaccine by the year 2007. Currently, however, there are only a small handful of phase III efficacy trials that have been initiated so far.

There are several major impediments to developing a successful vaccine for HIV-1. The first of these is the constant variability of the virus within the individual through constant mutation and recombination. Therefore, the significance of strain variation needs to be estimated in order to ensure the vaccine is capable of eliciting the correct response to the particular strain present. Secondly, the multiple virus subtypes make it difficult to design a vaccine that will target all of them. Investigations into the conserved regions of HIV genes to find proteins common to all or most subtypes will provide a better target, but if not found, then a vaccine 'triple cocktail' approach may be necessary. The easiest way to design an effective vaccine is to know what immune responses protect against it, and tailor a vaccine to turn on that specific response. This leads to the third problem as these so-called "correlates of immunity" have not being precisely identified for HIV and with no known documented case of complete recovery there remains the distinct possibility that a natural protective state does not exist. There are, however, HIV positive people who remain clinically asymptomatic over a long-term period and more importantly, people who have been multiply exposed but still remain uninfected. Studies on the immune response of these individuals could prove particularly useful in finding the correlates of immunity. Fourthly, as HIV can exist in the body as both a free virion and within an infected cell, any vaccine must activate both arms of the immune system. The first arm is humoral immunity whereby B-cells make antibodies to defend against the free virus. The other arm is cellular immunity that uses T helper cells (CD4+ T cells) and T killer cells (Cytotoxic T Lymphocytes, CTL's) to organise the immune response and attack infected cells. The most difficult challenge for a successful HIV vaccine is that the target of HIV is the immune system itself, in particular the T helper cells that instead of organising a defence are being modified or destroyed. As over 80% of transmission of the HIV virus arises sexually,[64] a

more daunting prospect is that a third type of immunity, mucosal immunity, may have to be stimulated. The immune cells lining the mucous membranes of the genital, respiratory, digestive, and reproductive tracts and those in the nearby lymph nodes are the body's first line of defence against infectious organisms. They produce different immune responses that are not well understood. The final hurdle to overcome is the lack of an ideal animal model in order to test vaccines against the introduction of a live virus.

The most extensively studied vaccine candidate is the recombinant gp120 subunit vaccine AIDSVAX developed by VaxGen.[65] They were the first to initiate a phase III trial for an AIDS vaccine in 1998 and followed this with a second phase III trial in Thailand in 1999.[66] Basically the recombinant gp120 (and also gp160) are used to evoke an immune response that will then trigger a larger attack on the free virus and infected cells.[65] Comparable to this method is the use of peptide epitopes that mimic only part of the HIV envelope instead of using the whole. However, peptide vaccines generally have several problems to overcome such as low immunogenicity, difficulty in translating results from animals to humans, extensive genetic variability of HIV, and lack of important structural conformation of some neutralisation epitopes of HIV. Another approach to developing vaccines is to use DNA expression plasmids that encode HIV-1 gene products that again, ultimately, are hoped to produce an immune response. After the first human trial[67] by the Apollon company of HIV Gag/Pol DNA vaccines, additional trials are underway. Other companies such as Merck, Chiron and Pasteur Merieux Connaught are also at various stages of developing similar DNA type vaccines. Along similar lines to the DNA vaccines are live vector vaccines. These are harmless bacteria or viruses containing genes that will encode for HIV proteins. These proteins are presented to the immune system, just as proteins from a virus-infected cell would be. Therefore, these vector-based vaccines produce both a humoral and cellular immune response and have the added advantage of being able to elicit a mucosal immune response as this is the route of their infection. Such vaccines based on the canarypox virus have entered human clinical trials.[68] The most controversial type of vaccine is the live-attenuated virus. While most licensed vaccines in use today are based on this concept there are serious safety issues with regards to HIV-1 due to the potential for the vaccine to cause AIDS instead of prevent it. However, research

is ongoing to improve the safety of these types of HIV vaccines but no human trials are planned as yet. A slightly safer option is to use whole, killed HIV that is incapable of replication but which will elicit the required immune response. In much the same way as anti-retrovirals, what may turn out to be the best approach is combination vaccine therapies which are being researched with some current phase II trials underway using the gp120 protein plus the canarypox vectors.[69] Using such combinations should hopefully stimulate both arms of the immune system and may have the benefit of not just having an additive but a synergistic effect. However, current approaches are mostly based on empiric observations and many complex clinical trial agreements would need to be established amongst the various manufacturers.

1.2 Cyanovirin-N

1.2.1 Discovery and Isolation

In 1992, Dr G. Patterson, University of Hawaii, as part of a larger NCI (National Cancer Institute) programme for the collection and screening of cyanobacteria, prepared an extract from a sample of the blue-green algae *Nostoc ellipsosporum*.[70] As part of the NCI's program to screen and evaluate natural product leads for anti-HIV activity, this extract was selected by the Laboratory for Drug Discovery, Research and Development (LDDRD) for further investigation due to its strong HIV inhibition and from preliminary indications that the active component was a protein.[71] Further anti-HIV testing and screening using an XTT assay showed significant activity that was not due to any known anti-HIV compounds such as sulfolipids or sulfated polysaccharides.[71] After dereplication and purification by reverse phase C18 HPLC an active protein was isolated.[71] The amino acid sequence of the protein was deduced from a combination of Edman degradation of both the intact protein and numerous overlapping fragments derived from endoproteinase digestion.[71, 72] The sequence obtained from these experiments matched an amino acid analysis of the protein.[71] Finally, mass spectrometry (ESIMS) showed a molecular ion in agreement with the calculated mass/charge ratio for the calculated molecular weight. The 101 amino acid protein with a molecular weight of 11,009 Da was named cyanovirin-N (CV-N) (**15**).[71] The sequence of CV-N (**15**) was compared to known proteins and gene products on various databases with no homologies beyond 8 continuous amino acids nor >20% total sequence homology for any known products. [71, 72]

1.2.2 Chemistry and Biochemical Testing

CV-N (**15**) is a particularly stable protein and can withstand multiple freeze-thaw cycles, organic solvents, 8M guanidine HCl, 0.5% SDS (detergent), 0.5% H_2O_2 or boiling (15min in H_2O) with no significant loss of anti-HIV activity.[71] There are four cysteine residues in

CV-N (**15**) involved in two cross-linked disulfide bonds and only when these are reduced and alkylated with vinyl pyridine is the activity eliminated.[71] One of the more exciting aspects of CV-N (**15**) is the range of its anti-retroviral activity. CV-N (**15**) showed potent inhibition with all tested laboratory strains of HIV-1, HIV-2 and SIV. Also, CV-N (**15**) showed analogous antiviral activity against T-tropic, M-tropic and dual tropic HIV-1 primary isolates.[71] The specificity of CV-N (**15**) was also demonstrated by its lack of activity with a series of unrelated human viruses such as herpesvirus type 1, cytomegalovirus and adenovirus type 5.[71]

In experiments with HIV-1_{RF} in CEM-SS cells it was found that CV-N (**15**) had a remarkable concentration-dependent inhibition of virus-induced cell killing that gave comparative decreases in the supernatant reverse transcriptase (RT) and viral core antigen (p24) levels.[71] In contrast, there was no observable toxicity or growth inhibition of control cells at even the highest concentrations of CV-N (**15**) used.[71] Also noted was the fact that for maximal anti-HIV activity, CV-N (**15**) had to be added before, or just after, addition of the virus. A delay of just 3 hours before addition of the CV-N (**15**) led to little or no activity. [71]

Further experiments with co-cultivation of both chronically infected and uninfected CEM-SS cells demonstrated that CV-N (**15**) was acting on both infected cell and virion. However, CV-N (**15**) was not preventing replication of the virus if a cell had already become infected prior to CV-N (**15**) dosing.[71] Boyd *et al.*[71] also showed that pre-incubation of uninfected cells with CV-N (**15**) followed by washing still led to viral infectivity pointing towards CV-N (**15**) acting as a virucide.[71] This means it was acting directly on some constituent of the virus and preventing it from infecting otherwise susceptible cells. This combined evidence indicated that the target of CV-N (**15**) might be the gp120 envelope protein of HIV exposed on both virion and infected cells. The timing suggested an interruption of the attachment and fusion process of both cell-cell and virus-cell transmission. To back up this hypothesis, CV-N (**15**) was shown to bind to free gp120 irreversibly to form some kind of complex, which in its entirety proved to have no anti-HIV activity.[71] Of the initial experiments performed with CV-N (**15**), the most interesting

were ELISA (Enzyme Linked Immunosorbent Assay) experiments demonstrating that the binding site of CV-N (**15**) to gp120 appeared distinct from that of monoclonal antibodies directed to the immunodominant V3 loop and, just as importantly, distinct from the CD4 binding site of gp120.[71] The failure of CV-N (**15**) to block gp120 binding to CD4, plus the previous experiments all pointed towards CV-N (**15**) acting after the initial binding of virus to cell, but before viral entry into the cell.

A study by Mariner *et al.*[73] in 1998 set out to investigate more fully the interactions between virus-associated gp120 and its relationship with CD4.[73] They demonstrated that CV-N (**15**) fails to block both soluble CD4 receptor (sCD4) binding to viral lysates (infected cells) and also binding of virions to cells. They concluded that CV-N (**15**) may bind to unique regions of gp120 which do not hinder its binding to CD4 receptors but prevent further fusion and viral entry.[73] However, the experiments did not prove that the binding activities measured, were in fact representative of the actual gp120-CD4 interactions. In an additional experiment, virions were treated to 100,000x EC_{50} of CV-N (**15**) with no observed change in virion morphological structure, unlike treatment with sCD4, which decreases the density of envelope spikes.[73]

Further work by Esser *et al.*[74] published a year later provided additional, more specific information about where CV-N (**15**) was (or was not) binding.[74] It was shown that CV-N (**15**) had no inhibitory effects on several monoclonal antibodies (MAb's) directed to the CD4 binding site on gp120.[74] Flow cytometry experiments indicated that CV-N (**15**) blocked binding of HIV virions at the Leu3A neutralising epitope on CD4, and furthermore blocked the Leu3A sensitive (CD4-dependent) component of virion binding to CD4+ cells.[74] These results seemed to imply that CV-N (**15**) was interfering with gp120-CD4 binding but this interpretation contrasts with another experiment by Esser *et al.*[74] that demonstrated that CV-N (**15**) did not impair sCD4 induced enhancement of exposure of certain gp120 epitopes. Therefore, according to this experiment, CV-N (**15**) did not affect sCD4 triggered conformational changes in gp120. The final experiment of this paper involved examining the effect of CV-N (**15**) treated virions on the glycosylation-dependent, neutralising epitope by MAb 2G12.[74] It was shown that CV-N (**15**) inhibited this epitope

region on gp120, which is characterised by high-mannose sugars. The 2G12 epitope overlies the stem of the V3 loop of gp120 and comprises domains from the C2, C3, C4 and V4 regions.[75, 76] It is also opposite, and 25Å away from the CD4 binding site. [76] The partial gp120 crystal structure indicates that there is a large, heavily glycosylated, immunologically silent region which may allow CV-N (**15**) to not only bind at or near the 2G12 site but at other sites in this region.[16, 44, 76]

After the conflicting results yielded by Mariner *et al.*[73] and Esser *et al.*[74], a simple direct biochemical analysis was carried out by Dey *et al.*[77] in an attempt to resolve the issue of what specific interactions were taking place. Specific binding of gp120 to CD4 on infected cells was followed and measured by Western blot assay techniques.[77] This directly demonstrated that CV-N (**15**) was indeed blocking the interaction between gp120 and CD4 in a dose-dependent fashion with no complications from possible co-receptor interaction, as the murine cells used did not possess them.[77] They also noted that CV-N (**15**) had been postulated to interfere with post binding activities but no experimental evidence had yet been put forward to establish this theory. Using cells that only expressed CCR5 but not CD4, it was shown that gp120 (preincubated with sCD4, necessary for binding) was strongly inhibited by CV-N (**15**).[77] Because of the possibility that CV-N (**15**) may directly interact with the CCR5 and CXCR4 receptors, an experiment was performed to determine this. No direct interaction was observed for CV-N (**15**) with either receptor, showing that the blocking effects of CV-N (**15**) are manifested only at the time of interaction with gp120 and the co-receptors.[77] To determine if the gp120-coreceptor binding inhibition was merely a consequence of CV-N (**15**) blocking sCD4 binding to gp120 (hence no activation or coreceptor interaction), an experiment was undertaken where CV-N (**15**) was added either before or after preincubation with sCD4.[77] The results indicated that there was no difference between the two protocols with respect to fusion inhibition. The only other possible explanation apart from direct CV-N (**15**) interference of gp120 coreceptor interaction, is that the sCD4 binding is reversible and in dynamic equilibrium. This would mean even though CV-N (**15**) was added after sCD4, it may bind to gp120 as the sCD4 dissociates, not allowing sCD4 to return, hence coreceptor blockage.[77] However, granting this possibility, in combination these findings point to CV-N (**15**) having two distinct

inhibitory effects in the HIV-1 Env fusion process. The first is direct inhibition of gp120 CD4 interaction post binding, and then further blocking of CD4 induced gp120 coreceptor interaction. Although previous studies had shown no other virus inhibition by CV-N (**15**) apart from HIV-1, HIV-2 and the closely related SIV,[71] it was thought with this new information about the dual blocking ability of CV-N (**15**), that further virus testing was warranted.[77] The feline immunodeficiency virus (FIV) expresses only the CXCR4 receptor (no CD4) but even at 10nM concentrations, CV-N (**15**) was able to strongly block infection.[77] A two further non-immunodeficiency type viruses, MV and HHV-6, that express CD46 as the entry receptor were also both inhibited by addition of CV-N (**15**).[77]

To add to this increasingly complex story O'Keefe *et al.*[78] showed how CV-N (**15**) bound to both sgp120 and sgp41 in a highly specific, fully saturable, glycosylation dependent manner. The interaction with gp41 is similar to that of gp120 in that the interactions are primarily electrostatic in nature and as both ΔG values are negative the binding appears spontaneous and favourable.[78] The binding constant for CV-N (**15**) to gp41 also appears to be at least a factor of 20 less than for gp120.[78] A further distinction is the stoichiometric ratios of the two interactions. It appears through both optical biosensor studies and isothermal calorimetric (ITC) experiments that approximately five molecules of CV-N (**15**) bind to each gp120 while ITC showed the relationship with gp41 is one to one.[78] No binding of CV-N (**15**) was detected when non-glycosylated gp120 or gp41 was used.[78] However, in binding strongly to CV-N (**15**), there was no distinction noted between using two very different glycosylated forms of gp120.[78] The recombinant baculovirus-produced sgp120 does not contain any *O*-linked oligosaccharides or any *N*-linked complex oligosaccharide chains found on the other H9 produced sgp120[79, 80] but contains only high-mannose type *N*-linked oligosaccharides.[81] Other glycoproteins tested such as α-acid glycoprotein, horseradish peroxidase, human serum albumin, and human IgGs showed no binding whatsoever to CV-N (**15**).[78]

The most recent paper by Mori and Boyd, published in 2001,[82] manages to answer some of the questions surrounding interactions of CV-N (**15**), gp120, and CD4. Their experiments involved coimmunoprecipitation analyses and flow cytometry-based sgp120-target cell-

binding assays.[82] The results conclusively indicated that CD4-dependent and CD4-independent binding of soluble gp120 (sgp120) to target cells were both strongly inhibited by CV-N (**15**) in a concentration dependent manner.[82] It was also found that more than one CV-N (**15**) molecule binding to sgp120 is required to inhibit both forms of sgp120 binding.[82] The second series of experiments indicated that CV-N (**15**) could block the sCD4-induced sgp120 binding to CXCR4-expressing cells.[82] There was no distinction between the blocking ability if CV-N (**15**) was added after or before addition of sCD4 to activate sgp120.[82] Therefore, it can be concluded that CV-N (**15**) leads to steric blocking and/or conformational changes to sgp120 regardless of whether sCD4 has induced a conformational change in sgp120 itself. As important as these first two series of experiments were, the third segment is undoubtedly the most intriguing result to come from this particular paper. CV-N (**15**) was found to induce sgp120 disassociation from target cells, again, from both CD4-dependent and CD4-independent binding.[82] There is no known antibody as yet that has been able to dissociate sgp120 from cell bound CD4. As Mab 2G12 is known neither to inhibit sgp120 binding or dissociation of bound sgp120 from CD4-expressing target cells the modes of action and/or binding sites of CV-N (**15**) and 2G12 must differ to some extent. This unique dissociative property of CV-N (**15**) was found to occur rapidly (less than 2.5 min) after sgp120-bound cells were incubated with CV-N (**15**). This could explain why CV-N (**15**) still showed activity even after a 60 min delay of addition of CV-N after exposure of cells to the virus.[71] The CV-N (**15**) could conceivably have been displacing the bound virion from CD4+ cells. Further work here would be vitally important in determining the implications and relevance of this novel property of CV-N (**15**).

A thorough explanation for these complex inhibitory effects of CV-N (**15**) is difficult to achieve as yet. Combining the results from Dey *et al.*[77] with those of Esser *et al.*[74] it is tempting to suggest that the diverse interactions of CV-N (**15**) are as a result of its affinity for carbohydrate moieties. In particular, the experiments by Esser *et al.*[74] in discovering how CV-N (**15**) blocks the 2G12 epitope on gp120 which is present in the immunologically 'silent face' lends support to this theory. The evidence for a direct interaction between carbohydrate functionalities and CV-N (**15**) was further enhanced with

the experiments by O'Keefe *et al.*[78] where CV-N (**15**) bound glycosylated, but not non-glycosylated, gp120 and gp41. However, this 'carbohydrate model' of interaction fails to explain how some distinctly different carbohydrate moieties bind strongly to CV-N (**15**) (HIV-1, HIV-2, SIV, FIV, MV and HHV6)[71, 77] while other glycosylated viruses and proteins have no affinity (herpesvirus type 1, cytomegalovirus, vaccina virus, human serum albumin, horseradish peroxidase, α-acid glycoprotein, human IgGs).[71, 77, 78] Compounding these results is the different range of receptors used by the viruses that do bind CV-N (**15**), from CD4 to CCR5 and CD46. The most recent results from Mori and Boyd[82] have shown that CV-N (**15**) binds to sgp120 in a way that alters or hides the 2G12 epitope, inhibits CD4-dependent and CD4-independent gp120 binding, and prevents sCD4-induced gp120 binding to cell-associated CXCR4. They finally showed how CV-N (**15**) has the remarkable ability to induce sgp120 dissociation from target cells which could explain several as yet unexplained results.[82] It now appears that CV-N (**15**) acts in three distinct stages of the Env fusion process. The primary site of inhibition appears to be gp120-CD4 interaction, with additional inhibition of CD4 induced coreceptor binding and finally a gp41 interaction that also presumably interferes with the conformational changes necessary for complete fusion. This three-pronged attack by CV-N (**15**) only serves to highlight again its remarkable properties of binding and anti-HIV activity that is unprecedented among other potential HIV therapies. More investigations into the specific binding positions and kinetics of these complex series of interactions are required before any conclusive theory can be formulated that explains all of the above data.

1.2.3 Recombinant Production

One of the advantages of isolating and determining the sequence of such a protein as CV-N (**15**) is that it is possible to 'grow' more of the compound through use of *Escherichia coli* (*E. coli*) expression techniques. [71] For CV-N (**15**), this firstly involved translating the amino acid sequence into a DNA sequence by selecting appropriate codons to represent each amino acid. Supplementary codons were included to facilitate incorporation of the gene into the expression vector, p-FLAG-1.[71] Transformation of *E. coli* was carried out

incorporating the synthetic gene sequence ligated into the *E. coli* genome. Induction of the clones produced an anti-HIV active FLAG-fusion protein that was purified by affinity chromatography.[71] Removal of the supplementary pieces of the coding by site directed mutagenesis left the final construct.[71] This was used to transform *E. coli* and CV-N-Cys was expressed from the recombinant organism.[71] Purification of the protein was by C4 reverse phase chromatography, followed by reverse phase C18 HPLC and finally C4 reverse phase HPLC with an increasing concentration of CH_3CN to H_2O.[71] Initially it was the HPLC retention time, ESI mass spectrometry and amino acid analysis that confirmed the identity and accuracy of the synthetic CV-N (**15**) compared to the natural protein. To provide further evidence, *N*-terminal sequencing was employed on the first 25 amino acids of each of the proteins and their anti-HIV activities tested together.[71] There was no difference observed between the two proteins throughout all the tests confirming the success of the transformation and expression process.[71]

In 1998 further work by Mori *et al.*[83] was published on this expression process for CV-N (**15**) and a modified approach was used utilising a different expression vector (pET-26b(+)). This was because of impure partial CV-N proteins present with the final product making purification difficult in the original process.[71] The revised method produced tight transcriptional control with homogeneous CV-N (**15**) and no sign of any truncated species.[83] The yield was also significant with approximately 10 mg of CV-N (**15**) per litre of cell culture.[83] The ability to produce large quantities of pure CV-N (**15**) is essential if it is to progress as an anti-HIV therapy of any kind. This is an advantage over other natural products that cannot be produced recombinantly and which are often difficult synthetic targets.

1.2.4 Sequence Requirements

To obtain further information about the binding characteristics of CV-N (**15**), a series of DNA constructs were produced and expressed in *E. coli* to create mutant forms of CV-N (**15**).[84] These mutants represented changes in the terminal makeup and functional domains

of the protein. A sequence comparison of the 101 amino acids in CV-N (**15**) yields some very intriguing information. The first 50 amino acids form Domain 1 (D1), and when compared to the next 51 amino acids of Domain 2 (D2), there is a striking similarity with 32% of the sequence identical, while a further 26% show merely conservative changes (**Figure 1.5**).[84] This gives a very high 58% overall homology that suggests within the cyanobacteria a tandem gene duplication of coding for CV-N (**15**).[84]

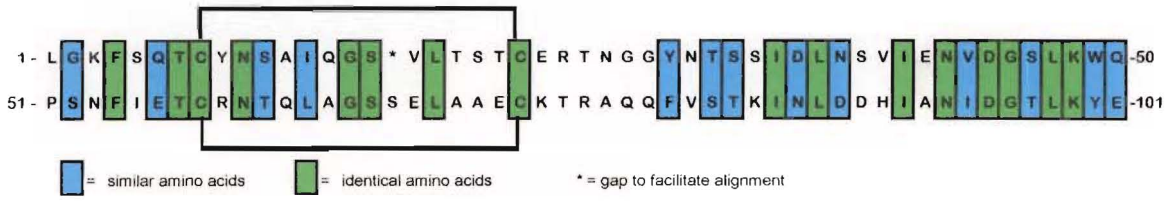


Figure 1.5 CV-N sequence comparison

The first series of mutants were constructed to study the effect of altering either the *N*- or *C*-terminus by deletion of amino acids. There were three *N*-terminal mutants, namely CV-N(-1N), CV-N(-2N) and CV-N(-3N) which removed 1, 2, and 3 amino acids respectively.[84] Removal of the single amino acid showed comparable anti-HIV activity to that of native CV-N (**15**) while the omission of 2 amino acids resulted in a 9-fold decrease in potency and the removal of the terminal tripeptide gave a 156-fold decrease in anti-HIV activity.[84] This trend was mirrored in the gp120 binding ELISA studies on the mutants which showed comparable gp120 binding of the CV-N(–1N) mutant with respect to native CV-N (**15**) while the (-2N) and (-3N) mutants showed only 73% and 50% gp120 binding affinity respectively.[84] The *C*-terminal mutants were made with an octapeptide leader (FLAG) to facilitate purification.[84] The control protein for this series was F-CV-N (F=FLAG), which although showing a modest decrease in potency compared to native CV-N (**15**), was sufficient to act as the standard.[84] Two mutants were constructed, F-CV-N(-3C) which lacked the *C*-terminal tripeptide and F-CV-N(-8C) in which the last 8 amino acids were deleted.[84] The F-CV-N(-3C) mutant showed a 42-fold decrease in anti-HIV activity and a 22% gp120 binding affinity compared to F-CV-N for each experiment.[84] The F-CV-N(-8C) construct however had no anti-HIV potency at all and essentially no gp120 binding affinity either.[84] These deletion mutants demonstrate how sensitive the respective

terminals of CV-N (**15**) are to modification upon their anti-HIV activity and gp120 binding affinities.

The second series of mutants involved swapping the domains of CV-N (**15**) around. The first of these was the circularly permuted F-D2D1 mutant, which showed relatively weak anti-HIV and gp120 binding activities.[84] The tandem mutants however, F-D1D1 and F-D2D2 did not show any activity for either assay.[84] The final series of mutants involved replacing the two pairs of cysteines of each disulfide bond with serine residues in turn, leading to the F-C>S#8-22 and F-C>S#58-73 mutants thereby lacking each of the disulfide bridges respectively.[84] Neither showed any anti-HIV activity but surprisingly the F-C>S#58-73 construct did show a modest gp120 affinity comparable to the F-D2D1 mutants.[84] While all of the mutants that displayed anti-HIV activity showed gp120 binding affinity, this final result would tend to suggest that while binding of gp120 is desirable, it is not the only requirement for blocking of HIV.

1.2.5 Solution and Crystal Structures

In 1998 the solution structure of CV-N (**15**) was solved by using double and triple resonance multidimensional heteronuclear NMR spectroscopy, utilising the uniformly ^{15}N- and $^{15}N/^{13}C$-labelled protein.[85] The final structural calculations were obtained by simulated annealing with 2,509 experimental NMR restraints including ^{1}H, ^{13}C, $3J$ couplings and residual dipolar couplings.[85] The dipolar couplings, unlike the other restraints which require a degree of spatial proximity of atoms (<5 Å), provide long-range structural information.[86] The shape of CV-N (**15**) in solution is an elongated prolate ellipsoid with a length of ~55 Å and a maximum width of ~25 Å. The secondary structure is comprised of 10 β-strands and four short 3_{10}-helical turns. As already shown in **Figure 1.5**, CV-N (**15**) can be easily divided into two series of sequence repeats, however, these repeat units do not form individual domains in the three dimensional structure (**Figure 1.6a)**). In reality, the overall folding pattern of the protein relies on multiple contacts between the two sequence repeats with this interaction burying 3,085 $Å^2$ of accessible surface area.[85]

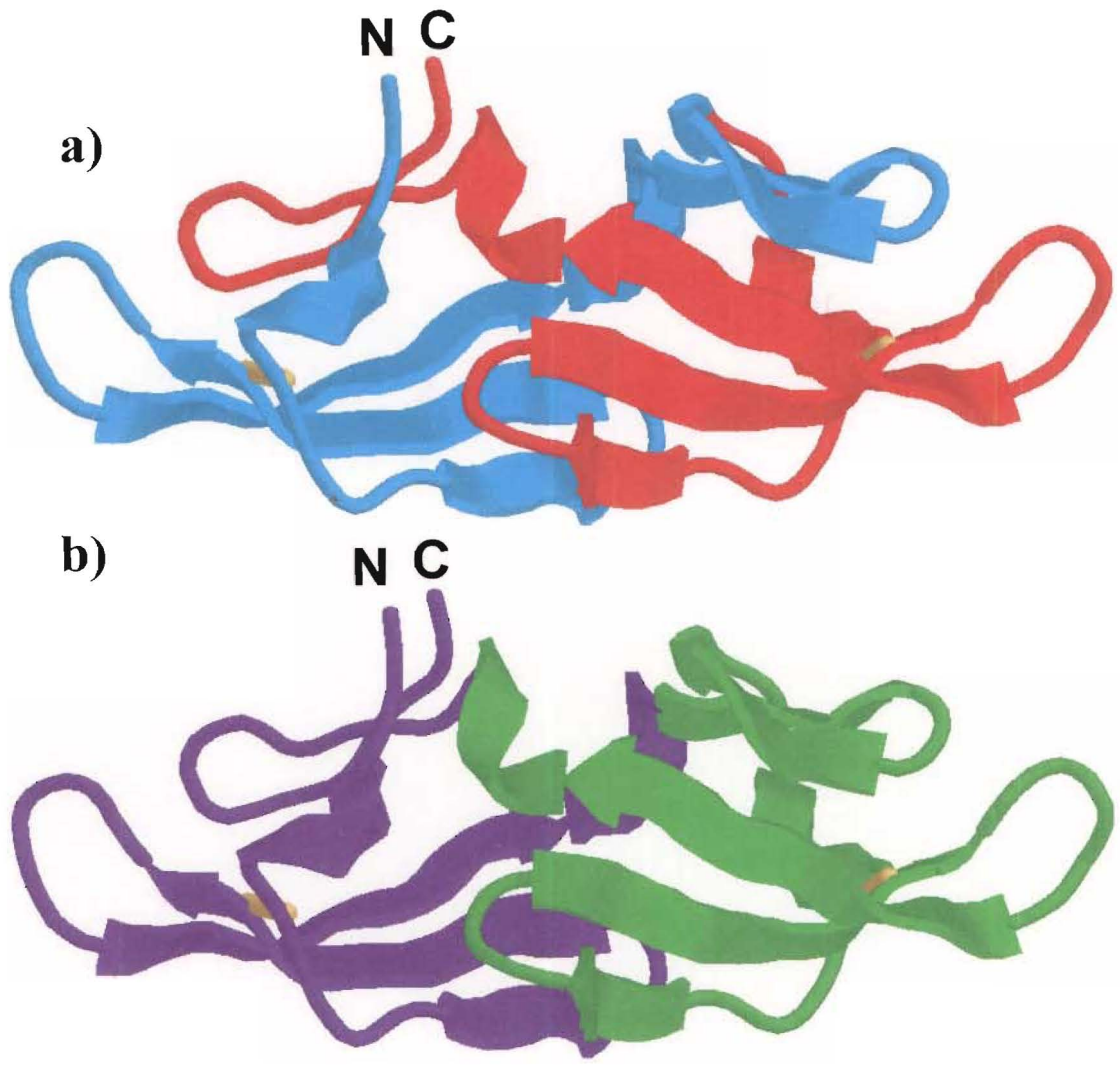


Figure 1.6 Internal 2-fold symmetry of CV-N
a) Ribbon diagram with two sequential repeats coloured blue (residues 1-50) and red (residues 51-101)
b) Ribbon diagram with structural domains A coloured purple, (residues 1-39 and 90-101) and B coloured green (residues 39-90)

The structure can be divided into two symmetrically related domains, A and B (structure based domains), which are formed by a strand exchange between the two sequence repeats (**Figure 1.6b**)).[85] Both *N*- and *C*-termini are contained within domain A which comprises residues 1-39 and 90-101, with domain B containing the residues 39-90. As viewed in **Figure 1.6b)**, these represent the green and purple sections respectively which will be referred to as 'structural domains'. These domains can be overlaid to produce a Cα atomic

r.m.s. difference of only 1.3 Å.[85] One of the most striking features that is a product of the internal two fold pseudosymmetry is that the *N*- and *C*-termini are adjacent to each other with van der Waals interactions between the side chains of Leu1 and Glu101.[85] The closeness of these residues in space would tend to support and explain the mutagenesis studies of Mori *et al.*[84] as deletion of amino acids from either terminus resulted in a loss of anti-HIV activity. Presumably the removal of the terminal peptide regions would disrupt the secondary structure with loss of the electrostatic 'bridge' joining residues 1 to 101.

When proteins interact with each other (as with CV-N (**15**) and gp120) the surface hydrophobicity of each protein plays a crucial role in determining this interaction.[87-90] Therefore, Bewley *et al.*[85] mapped the most hydrophobic regions of CV-N (**15**) using the method of Covell and co-workers[88-90] in order to predict regions which may interact with gp120. For example, the most hydrophobic cluster involves residues Leu1, Gly2, Lys3, Gln6, Thr7, Thr25, Asn26, Asn93, Ile94 and Asp95, and is centered around the *N*-terminus.[85] This also helps explain how the gp120 binding and anti-HIV activity are reduced to nothing with loss of the *N*-terminal tripeptide.[84] Other hydrophobic clusters can also be mapped and with further structural relationship studies, the interaction between CV-N (**15**) and gp120 can be more closely defined.

A year later in 1999 the crystal structure of CV-N (**15**) was solved by Yang *et al.*[91] which allowed a direct comparison of the two structures and also a direct comparison of the accuracy of each method. Using the NMR structure as a model, a solution with the correct space group enantiomorph was quickly found and the structure well defined except for one region, residues 50 to 53.[91] This region links the two sequence based domains and in the NMR structure showed no distortions. To attempt to overcome this disparity a model was built that incorporated two domains from two monomers connected by a new linker.[91] This domain-swapped (dimeric) form reduced the uncertainty considerably giving an acceptable quality structure.[91] In solution, CV-N (**15**) primarily exists as a monomer and so the discovery of this domain swapping was unexpected. New crystals were grown from a sample verified by gel filtration and NMR to be monomer but after identical refinement the domain-swapped structure was again resolved.[91] Investigations into the conditions

necessary for formation of dimeric domain-swapped CV-N revealed that reverse-phase HPLC under acidic conditions (0.05% TFA) gave small amounts of dimeric product and obviously crystallisation conditions (low pH and 26% isopropanol) also allowed formation of dimer.[91]

The linker region from the Cα atoms of residues 49 to 54 is extended in the domain-swapped structure to 13.8 Å from 10.2 Å in the solution structure.[91] Apart from this linker region and a few flexible side chains, there was excellent compatibility between the two structures and a very high degree of similarity between the domains.[91] It is important to keep in mind that without the NMR solution structure for comparison, the resolved crystal structure would have given the protein a form that is not the biologically significant one. There is no evidence as yet to suggest any properties of the dimer that may contribute to CV-N's (**15**) anti-HIV activity but the possibility does exist and should be noted for further studies.

An addition to this structural work was made by Bewley *et al.*[92] with the publication of the dimeric form of CV-N in solution. Whereas the X-ray dimer can be considered as two CV-N molecules side by side (albeit with units AB' and A'B), the solution structure was resolved to show the two halves of the dimer anti-parallel with each other.[92] The reasons for these differing structures have been postulated by Bewley *et al.*[92] The carboxylates of Glu41 and Glu41' hydrogen bond in the crystal structure but in solution at neutral pH they would be repelled away from each other. The X-ray structure also shows the stabilising effect of a neighbouring subunit inserted between the two halves of the dimer. The absence of crystal packing in solution and the lack of hydrogen bonds between the glutamic acid residues is believed to lead to a re-orientation of the two halves of the dimer.[92]

1.2.6 Applications of CV-N

1.2.6.1 Microbicide

From an early stage, due to its characteristic properties, CV-N (**15**) was investigated as a possible antiviral microbicide, which are products that can be used topically (*ex vivo*) to prevent the spread of the virus. There continues to be a enormous need for a female controlled topical microbicide that can block HIV infection and the World Health Organisation, the National Institute of Allergy and Infectious Diseases, the U.S. Department of Health and Human Services and many others have acknowledged this form of prevention as a critical global priority.[93-96] CV-N (**15**) is a potential candidate for this type of application, particularly as any microbicide of this form would have to inhibit free viral particles and infected cells which are both transmitted sexually. There is also the added advantage that CV-N (**15**) is active against both M-tropic (macrophage-tropic) and T-tropic (CD4 T-cell-tropic) primary isolates of HIV-1 with significant evidence that M-tropic forms of the virus are crucial for sexual transmission.[41, 97-99]

1.2.6.2 CV-N-PE38 Construct

Because of CV-N (**15**)'s unique ability to selectively target and bind to gp120 across various M- and T-tropic isolates, as well as the virion itself, there is the potential for using CV-N (**15**) as a targeting moiety with suitable compounds attached that will actively kill infected cells and free virus. Similar compounds have been made that use sCD4 as the targeting moiety but although early results were promising *in vitro*, the gp120 targeting ability in fresh clinical isolates and HIV+ patients was dramatically reduced.[100-107] With the potent activity of CV-N (**15**) shown against a variety of clinical isolates in mind, a chimeric toxin molecule was constructed consisting of CV-N (**15**) and the potent, active section of *Pseudomonas* exotoxin (PE).[108] This was done by constructing an *E. coli* expression plasmid that contained a gene coding for both FLAG-CV-N and amino acids 253-364 and 381-613 of PE.[108] This section of PE, labelled PE38, contains domains responsible for translocation and ADP-ribosylation.[109, 110] Purification by affinity

chromatography and gel filtration led to the single product of approximate molecular mass 49.3 kDa.[108] In subsequent, testing the toxin conjugate did show enhanced specificity for killing of chronically infected gp120 expressing H9/HIV$_{IIIb}$ cells with a much lower IC_{50} (0.004 nM) than for the control, healthy, non-gp120 expressing cells (IC_{50}=0.3 nM).[108] Another experiment tested the competition between F-CV-N-PE38 and native CV-N (**15**) for H9/HIV$_{IIIb}$ cells where addition of the native protein reduced the toxicity of the conjugate to that of uninfected H9 cells.[108] An ELISA binding assay with gp120 treated plates showed that the F-CV-N-PE38 had distinctly less binding affinity for gp120 than native CV-N (**15**).[108] This reduced affinity can be somewhat expected with the ~38 kDa addition of a large protein mass to the 11 kDa CV-N (**15**) molecule sterically hindering the interaction between gp120 and CV-N (**15**). Although the binding was somewhat reduced, this compound does show the potential for such an application.

1.2.6.3 Immobilised CV-N

Another novel approach to the application of CV-N (**15**) is in solid phase immobilisation in order to screen out HIV in blood and plasma products. Gandhi *et al.*[111] achieved this by biotinylating CV-N (**15**) and coupling this to streptavidin-coated glass porous magnetic beads (SGPM-beads). Tests on biotinylated CV-N (bCV-N) against native recombinant CV-N (**15**) in an anti-HIV assay showed that the biotin did not interfere at all with the anti-HIV activity of CV-N (**15**).[111] More importantly the sessile CV-N (sCV-N) also showed potent inactivating ability for the two HIV strains tested.[111] However, RT-PCR tests on the virus attached to the sCV-N and on the supernatant left behind showed that only a fraction of the virus was removed by the sCV-N with the remainder of supernatant being non-infectious virus (replication incompetent virions (RIV)).[111] This RIV is being investigated as a candidate vaccine due to its non-infectious, but still intact, whole virion status. Tests were also carried out under a variety of conditions to ensure that the CV-N that was bound to the beads did not leach out into the supernatant and all results indicated that this indeed did not occur.[111]

1.2.6.4 CV-N Based Assay

A recent application of CV-N (**15**) takes advantage of its specific, tight binding to unique regions of gp120 in order to screen thousands of natural product extracts for binding at, or near, the site(s) for CV-N (**15**).[112] Previous assays based on soluble forms of gp120 and CD4 have been used in a similar way in order to screen large collections for inhibition of their particular interactions.[113] The assay system used for CV-N (**15**) was dissociation-enhanced lanthanide fluoroimmunoassay (DEFIA®, PerkinElmer Wallac Inc., Gaithersburg, MD, U.S.A.) which is well suited to screening crude natural products. The process involved pre-incubating the extracts in wells with gp120, adding europium labelled CV-N, incubating, washing, and detection by release of the Eu^{3+} in a fluorescent form with an enhancer solution.[112] An assay 'hit' was defined as an extract that inhibited the binding of Eu^{3+}-CV-N to gp120 by 60% or more.[112] Of the 50,865 extracts tested, covering a range of terrestrial plants, marine algae, microbes, marine invertebrates and dinoflagellates, there were a total of 120 'hits' (64 organic and 56 aqueous).[112] This corresponded to a very low average hit rate of 0.2%, which is possibly due to CV-N's (**15**) very high affinity for gp120 and/or the unique binding site utilised by CV-N (**15**).[112] None of the organic 'hits' were able to be reproduced in subsequent assays but 51 of the 56 aqueous extracts were confirmed to be positive.[112] These 51 samples were then further tested for their ability to bind to C4 in a chromatographic column as previous work by the NCI had shown that aqueous extracts often contained high molecular weight polysaccharides and polyanionic compounds that non-specifically inhibited HIV infection[114] and did not bind well to C4. Of the 51 C4 retentates, 29 continued to inhibit Eu^{3+}-CV-N binding to gp120, of which 7 had significant *in vitro* anti-HIV activity.[112] However, in each case the activity was subsequently determined to be due to polyphenolic tannins, which are not considered suitable for drug development.[112]

1.3 Natural Products

1.3.1 Marine Natural Products

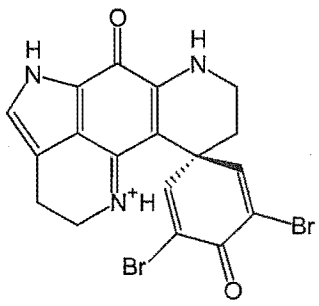
Natural products, the secondary metabolites produced by living organisms, have been extracted by Man and exploited throughout the ages for an extensive range of applications, from food to fragrances, pigments to poisons, and microbicides to medicines. Investigations into organisms of terrestrial origin has obviously been much more extensive as access and familiarity with the land have provided the 'path of least resistance'. However, as two third's of the surface of Earth is covered in oceans, this makes for a vast array of biodiversity in an estimated 1.5-4.5 million species that has only relatively recently begun to be extensively studied. The last three decades have been marked by intense interest in the chemicals in marine organisms and their biological activity. This rise in interest can be significantly traced to the availability of Self Contained Underwater Breathing Apparatus (SCUBA), which has allowed a much further exploration of the oceans to depths of up to 40 m while mini-submersibles allow collection down to depths of 1,000 m or greater.

The saline environment seems to encourage the production of novel structures due to the high levels of halogens and nitrogen, as well as the need for soft sessile organisms to chemically defend themselves from predation. Indeed, it has been shown that many marine organisms produce natural products while under stress[115] and furthermore that natural products are common in organisms without an immune system.[116] Both these studies highlight the role of natural products as a chemical defence system. An interesting feature of marine natural products research is the variability in the chemicals found in an organism depending on the location and season. There is mounting evidence that this is perhaps tied to the role of symbiotic microorganisms in the production of the chemicals.

1.3.2 The Marine Chemistry Group

The Marine Chemistry Group at the University of Canterbury began an intensive collection program in 1983 to exploit the high levels of bioactivity found throughout marine invertebrates. Over the next few years, collections were made from such diverse places as McMurdo Sound in the Antarctic, to the sub-tropical waters off the coast of northern New Zealand and throughout waters of both North and South Islands. Small-scale extractions were carried out on each sample to determine their biological activity and subsequent dereplication of larger extractions undertaken in order to elucidate the active component(s) to assess their pharmaceutical potential.

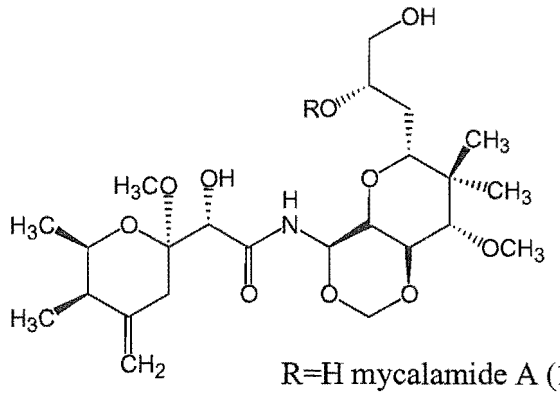
Several species in particular, have demonstrated potent activity in the in-house assay system and have therefore been extensively studied. The first of these is sponges of the genus *Latrunculia* (Order Hadromerida) found in waters ranging from Antarctica to NE New Zealand. They have produced an entire family of biologically active compounds called the discorhabdins. The compounds are readily identified by their red or green colour due to the 1,3,4,5-tetrahydropyrrolo[4,3,2-*de*]quinoline ring system. Discorhabdin C (**16**) is the simplest of the family but all are strongly cytotoxic against the P388 and BSC cell lines. Structural Activity Relationships (SAR's) have been carried out, particularly with discorhabdin C (**16**) in order to determine the functional source of the biological activity.[117]



discorhabdin C (**16**)

Sponges of the genus *Mycale* (Order Poecilosclerida) have produced one of the most potent antitumour and antiviral compounds to be isolated from a New Zealand programme. Purification of the active crude extract using bioassay guided fractionation led to the isolation of not one but two compounds, mycalamides A (**17**) and B (**18**).[118] These compounds showed antiviral activity but proved to be too cytotoxic to be antiviral agents.

Similarly they were also rejected as antitumour agents after *in vivo* trials highlighted problems with cytotoxicity.



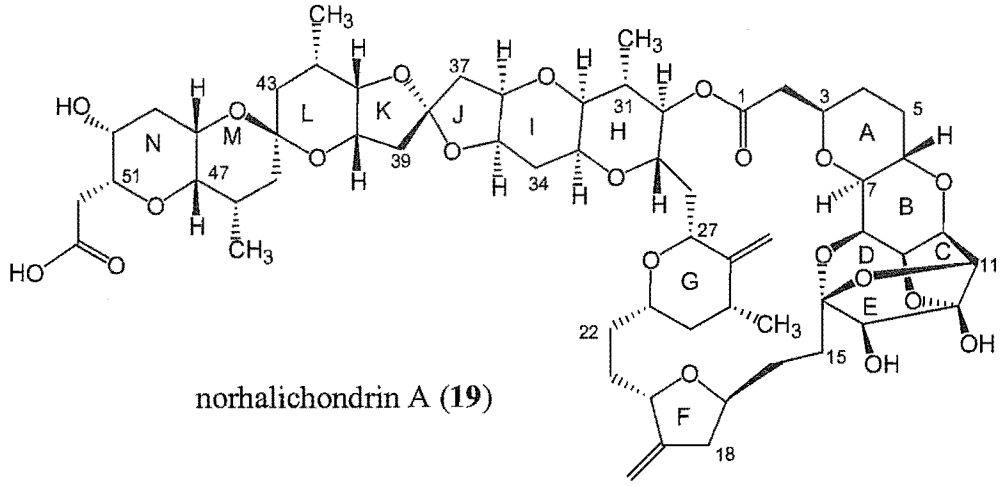
R=H mycalamide A (**17**)

R=CH_3 mycalamide B (**18**)

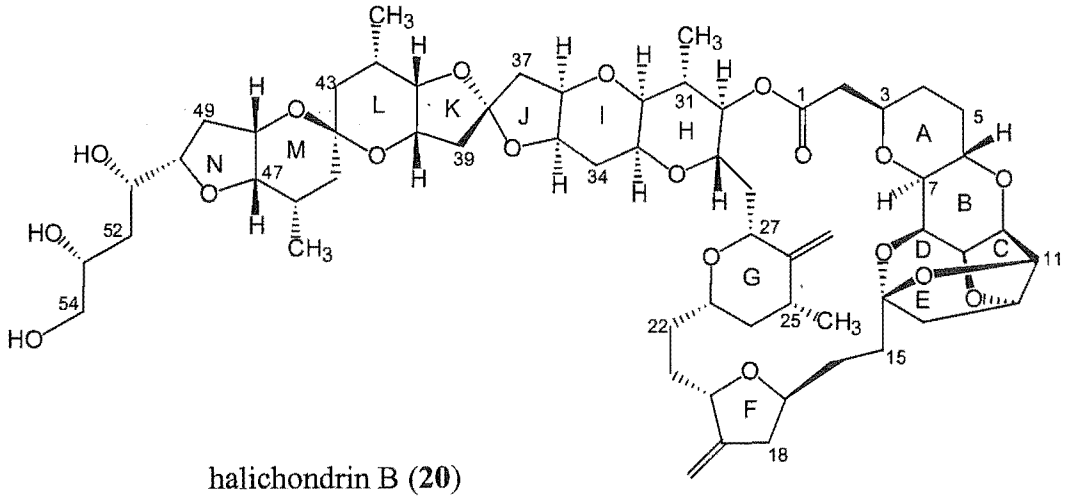
The Marine Chemistry Group at the University of Canterbury continues to search and isolate novel bioactive compounds for their pharmaceutical potential. However, the dramatic increase and interest shown around the world in the last decades for isolating natural products has meant that the probability of finding such novel compounds has decreased significantly. Recently, the Group has branched out to investigate marine fungi for bioactive compounds, which may hold as yet untapped sources of novel natural products. Although worldwide the number of new compounds isolated each year has stagnated recently, there remains a vast undersea resource to be explored with more advanced collection techniques utilising new technologies. The analogy of a treasure hunt is an apt one albeit with the caveat that participants must draw the treasure map as they go.

1.3.3 The Halichondrins

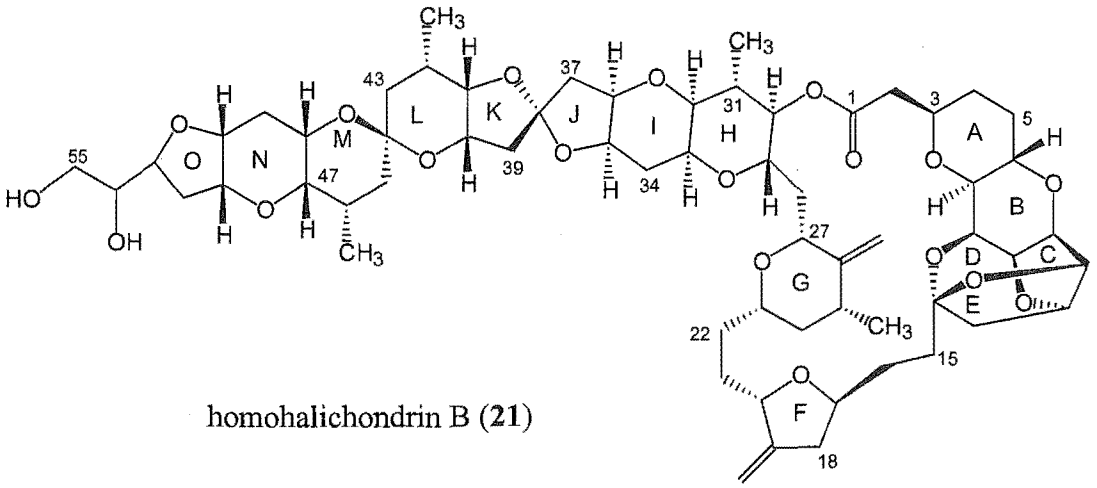
The series of compounds known as the halichondrins were first isolated from the Japanese sponge *Halichondria okadai* Kadota in the mid 1980's.[119, 120] These compounds, norhalichondrin A (**19**), B and C, halichondrin B (**20**) and C, and homohalichondrin A, B (**21**) and C represent three classes of the halichondrin family. These A, B, and C families are distinguished by the degree of oxidation at C12 and C13 (**Figure 1.7**). Within the families the differences lie beyond the C45 position. The structures are characterised by a novel 2,6,9-trioxatricyclo[3.3.2.0]decane system (rings C-E), a 22-membered lactone ring (C1-C30), two exocyclic olefinic groups and several pyranose and furanose rings.



norhalichondrin A (**19**)



halichondrin B (**20**)



homohalichondrin B (**21**)

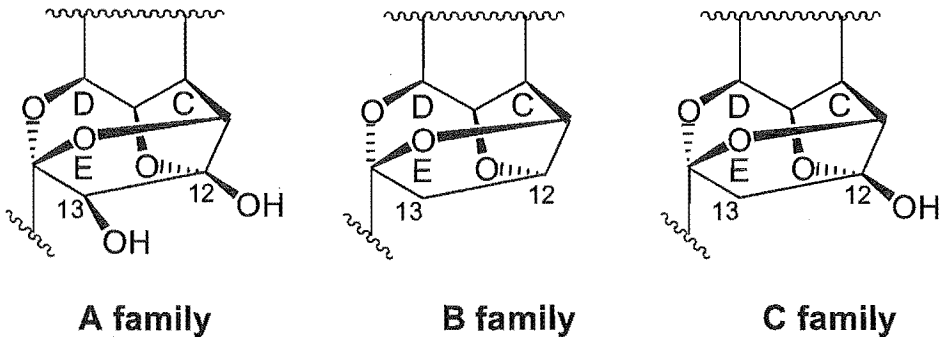
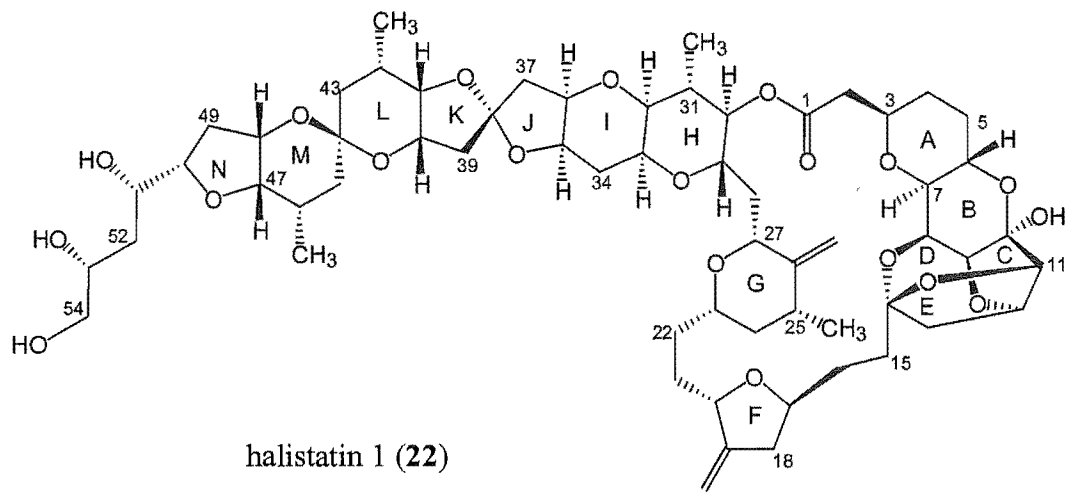
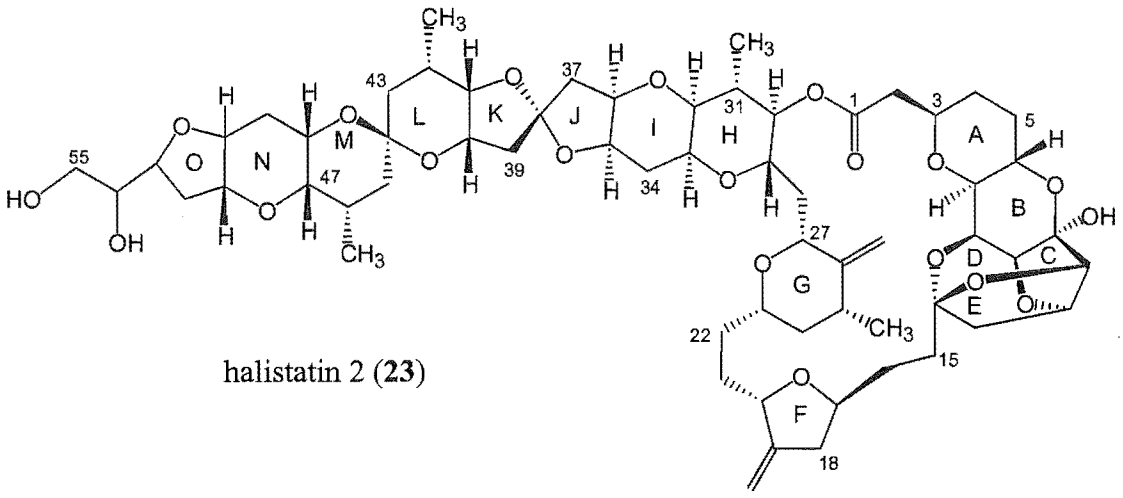


Figure 1.7 Halichondrin A, B, and C families

In 1991 halichondrin B (**20**) and homohalichondrin B (**21**) were isolated from an unrelated sponge *Axinella* sp., collected in Palau.[121] Two years later Pettit *et al.* isolated two new halichondrin compounds, halistatin 1[122] (**22**) and halistatin 2[123] (**23**). The sponge that halistatin 1 (**22**) was isolated from, *Phakellia carteri*, was also found to contain halichondrin B (**20**) and homohalichondrin B (**21**). To add to this complex series of interconnections, all three of these compounds were also isolated from a Western Indian Ocean sponge *Axinella* cf. *carteri* Dendy together with halistatin 2 (**23**).

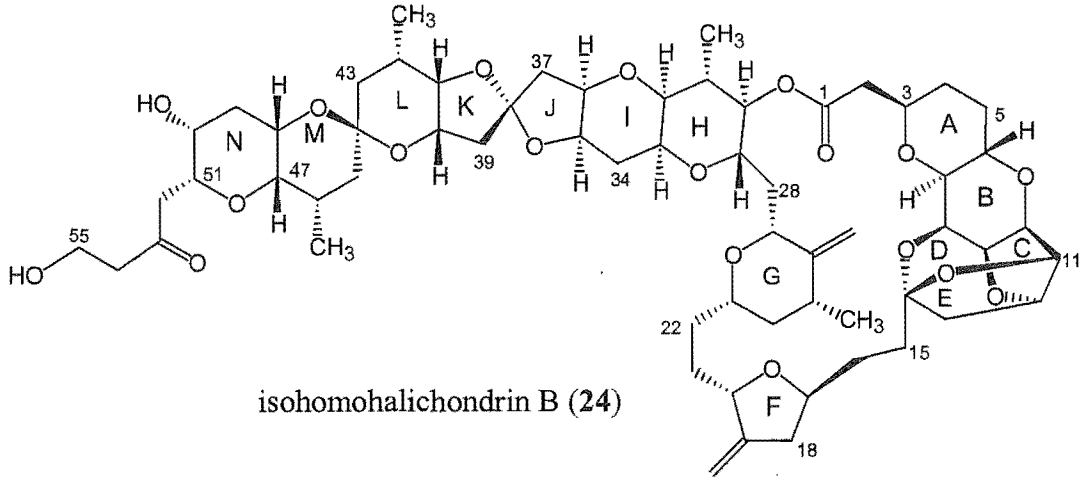


halistatin 1 (**22**)



halistatin 2 (**23**)

The Marine Chemistry Group at the University of Canterbury joined this story in 1987 when Rob Lake identified halichondrin-like compounds from two unrelated species of sponge. These sponges were *Raspalia agiminata* (order Axinellida, family Raspailiidae) from the Leigh area of the North Island[124], and the second, a slimy yellow sponge off the Kaikoura coast identified as one or more new species of the genus *Lissodendoryx* Topsent (class Demospongiae, order Poecilosclerida, family Myxillidae).[125] An interesting aspect of the Kaikoura sponge extraction was that it was the highest yielding (~1 mg total halichondrins/kg wet weight of sponge) of any of the other halichondrin producing sponges.[119, 121, 122] A subsequent, larger extraction carried out by Mark Litaudon of the *Lissodendoryx* sp. yielded 4.5 mg of an unidentified halichondrin. Characterisation utilising 2D-NMR and MS techniques yielded the compound isohomohalichondrin B (**24**)



isohomohalichondrin B (**24**)

which was named because it is isobaric to homohalichondrin B (**21**). The difference between them exists beyond C48, where isohomohalichondrin B (**24**) has the ring opened form of the terminal furanose unit. The activities of some of the halichondrin compounds are shown in **Table 1.1**. The three assays are cytotoxicity against B16 Melanoma cells, the Marine Chemistry Group's in-house P388 *in vitro* murine leukemia assay, and against the NCI's 60 cell-line *in vitro* primary screen.

Table 1.1 *In Vitro* Cytotoxicities of selected Halichondrins

Compound	B-16 IC_{50} (ng/mL)	P388 IC_{50} (ng/mL)	NCI GI_{50} ($x10^{-10}M$)
halichondrin B (**20**)	0.093	0.78	1.38
homohalichondrin B (**21**)	0.1	0.22	3.16
isohomohalichondrin B (**24**)	not recorded	0.18	1.15

Evaluation of the pattern of differential activity (mean-graph profile) of halichondrin B (**20**) against the NCI 60 human tumour cell line panel using the COMPARE[126] algorithm found that it most closely resembled those of tubulin binders such as vincristine and taxol. This strongly suggested halichondrin B (**20**) was an antimitotic agent. A significant feature of cancer cells is that they divide more rapidly than healthy cells. As microtubules are involved in cell proliferation (mitosis), microtubule inhibitors such as halichondrins are useful tools to treat cancer. Studies by Bai *et al.*[127] indicated that halichondrin B (**20**) was binding in the vinca domain of the microtubules. These vinca domain drugs inhibit tubulin-dependent GTP hydrolysis and also inhibit GTP/GDP exchange at the exchangeable nucleotide site of tubulin.

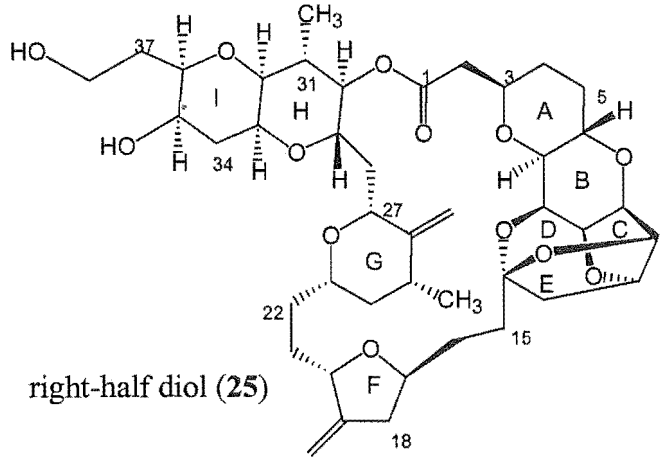
The most potent member of the family, halichondrin B, has undergone further testing at the NCI. The sixty cell line *in vitro* cancer screen run by the Developmental Therapeutics Program at the NCI demonstrated the potency and differential activity of halichondrin B (**20**) against several human tumour xenograft models. The compound has reached Decision Network IIA status but is unlikely to progress further until a sufficient supply is made available. The three alternative supply routes appear to be sponge collection,

aquaculture or synthetic manufacture. There are several problems with each of the above methods which make supplying the necessary amounts required an onerous task.

The natural supply of the highest yielding of the halichondrin containing sponges *Lissodendoryx* sp., found exclusively off the Kaikoura peninsula, is limited. The supply has been estimated using an ROV and benthic camera to be 289 ± 90 tonnes which is insufficient for an estimated worldwide demand of 5,000 tonnes to yield 5 kg a year of halichondrin B (**20**).

The second option of aquaculture has shown some initial encouraging results.[128-130] However, the levels of halichondrin-like compounds in the aquacultured samples were lower when compared to the native sponges, but there is potential for optimisation of conditions.[130] It was observed that the highest rate of halichondrin production was in sponges suffering from stress which correlates to the theory of sponges using the toxins as a defence mechanism.[130]

Because of the unique nature and challenging structure of the halichondrins, much effort has been put in to their synthesis over the years.[131] Halichondrin B (**20**) and norhalichondrin B were completely synthesised by Kishi and coworkers in 1991.[132] However, the usefulness of this impressive accomplishment is tempered by the fact that it took approximately 120 steps to achieve. Since then several improvements have been made to this synthesis.[133] This has included the complete synthesis of the right-half diol (**25**) which has shown the same activity pattern as the parent halichondrin B (**20**) with a cytotoxicity within an order of magnitude of the parent.[134] This offers the potential for a commercially viable synthesis of a halichondrin based compound at much reduced effort compared to the full structure.



right-half diol (**25**)

Future production of halichondrin compounds may also be possible if the organism producing it could be identified and the genetic component corresponding to halichondrin production spliced into an organism capable of fermentation. However, it is still unclear as to whether it is the sponge cells or a symbiotic organism that produces the toxin. Several groups are working on this problem while others are attempting to propagate sponge cells in culture which may provide a further alternative source but these studies are only at preliminary stages.

1.4 Project Aims

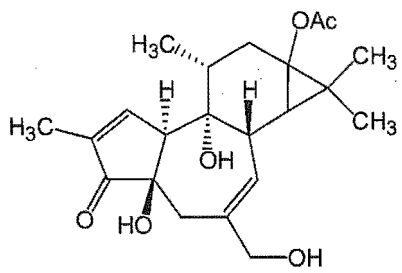
The overall aim of this project was to develop and act upon a strategy to produce CV-N-toxin conjugates capable of eradicating HIV-infected cells. This work primarily carries on from the experiments of Mori *et al.*[108] outlined in section **1.2.6.2** where CV-N was produced in tandem with a fragment of *Pseudomonas* exotoxin A attached. The conclusions reached indicated that although cytotoxicity was enhanced against gp120 expressing cells, the size of the PE38 fragment prevented effective gp120 binding therefore limiting the usefulness of the construct.[108] What this thesis sets out is a synthetic strategy to make CV-N constructs containing much smaller toxins making them capable of binding efficiently to gp120. These compounds therefore optimise selectivity of killing of HIV-infected, *versus* uninfected cells.

One of the crucial aspects of this strategy will be the internalisation of the constructs once formed. The gp120 is expressed on the outside of infected cells and most virologists consider that the envelope proteins are in the process of leaving the cell *via* viral budding. However, the efficacy of toxin conjugates that target the envelope proteins is such, that at least a portion of the gp120/gp41 complex is reinternalised.[135] Studies by Pincus[135] have demonstrated that not only is there a recirculating pool of envelope protein, but that the rates of internalisation are subject to regulatory influences.

As noted (section **1.2.6.2**) there have been several HIV-targeted toxin conjugates made in different labs utilising a range of toxins and carriers.[136-138] These include using free CD4 and IL-2 (IL-2 receptors are found on activated T-cells) as targeting moieties and toxins that include ricin A chain and truncated forms of diptheria toxin and *Pseudomonas* exotoxin A. The CD4-*Pseudomonas* exotoxin performed wonderfully in *in vitro* studies, but *in vivo* clinical trials raised serious issues of hepatotoxicity.[106, 107] However, Berger *et al.*[139] published a discussion of the reasons for this disappointing result in 1998 which offered renewed hope for this approach with insights that may be applicable to the results of this study. At present an IL-2-diptheria toxin fusion is also undergoing clinical trials.

Although the disappointing results of the CD4-pseudomonas exotoxin conjugate markedly diminished the enthusiasm for the development and clinical testing of similar constructs, the concept is valid and only awaits a more suitable candidate.

Another issue that would need to be addressed is the pool of latently infected cells (section **1.1.7**) that have the potential to 're-initiate' active virus production after levels had been reduced by anti-retroviral therapies. To be truly effective at eliminating all HIV from the body whether latently or actively infected cells, the type of CV-N-toxin conjugates under investigation would need to be introduced with compounds that 'turn on' the latent reservoir. There are several compounds that have this remarkable property, one of which, prostratin (**26**), has been proposed for this purpose.[140] The idea would be to initiate virus production in the latently infected cells with prostratin (**26**), therefore 'identifying' the cells to the CV-N-toxin conjugates which would then target and destroy them. The aim would ultimately be complete eradication of the virus from HIV infected individuals, something that has not yet been clinically achieved.



prostratin (**26**)

The limitations imposed on the synthesis of the toxin conjugates in this project were as follows.

- The alterations made to the native CV-N (**15**) must not interfere significantly with the biological binding profile of the CV-N (**15**) to gp120.
- Any synthetic modifications to the toxins must not inhibit their toxicity.
- Ideally the toxin should not become 'active' until internalised within the infected cells.
- Synthetic steps must ideally be high yielding particularly when dealing with toxins.

Overall this work was undertaken to combine the remarkable binding and stability of CV-N (**15**) with the potent toxicity of marine natural products such as the halichondrins. The fusion of these two distinct biological profiles was aimed at providing a specific yet powerful cell killing strategy.

Chapter 2

EARLY APPROACHES AND STRATEGY DEVELOPMENT

2.1 Introduction

Early work for this project focused on the best strategies to couple together CV-N (**15**) with the toxins available. The starting point was the production of CV-N(-3N) by Mori *et al.*[84] mentioned in section **1.2.4**. This suggested a method to synthetically produce the 'missing' terminal tripeptide with toxin attached, then couple to the truncated CV-N(-3N) and hence produce an intact CV-N-toxin conjugate.

As more aspects of protein modification strategies were investigated, more useful and productive routes were developed incorporating cleavable biolinkers. These biolinkers would be the basis by which the toxin would be released *in vitro* and *in vivo* once inside an HIV-infected cell.

Other methods involved activation strategies in order to promote the coupling of CV-N (**15**) functionalities with modified toxin constructs. Two activation methods stood out as suitable with one relying on reaction of CV-N's (**15**) lysine side chain amines with succinimidyl esters and the other utilising thiols added to CV-N (**15**) in order to react with maleimide functionalities.

The development of suitable toxin derivatives capable of surviving various chemical reaction conditions and synthetic steps was crucial as the toxins represent the killing power of the final constructs. A careful selection was made from the group of toxins isolated by

the Marine Chemistry Group over the past decades. The halichondrin derivatives were chosen for their potent activity, the extensive in-house knowledge of their chemical behaviour and ease of modification.

2.2 Terminal Tripeptide

The initial strategy developed to synthesise CV-N-toxin conjugates was based on the work by Mori *et al.* at the LDDRD in recombinantly producing CV-N(-3N).[84] As the 'missing' terminal tripeptide represents the residues Leu-Gly-Lys, there was potential to attach the toxin molecule through the side chain of the lysine residue. This method had the advantage of minimal modification to CV-N (**15**) as once the tripeptide was 'reunited' with the CV-N(-3N) the resulting product should be intact CV-N with a toxin attached in one position only (**Figure 2.1**).

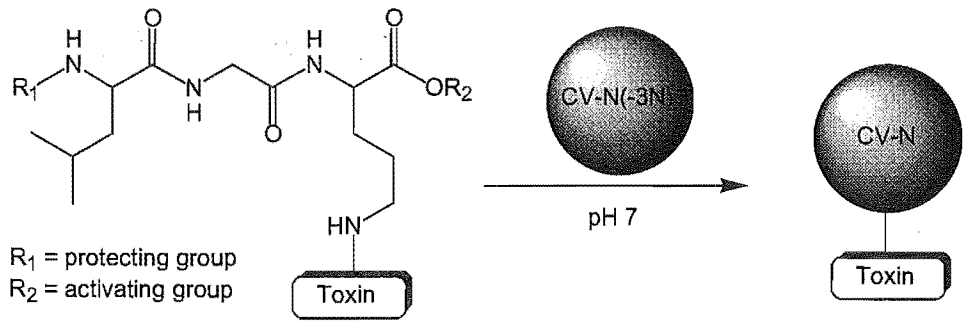


Figure 2.1 *N*-terminal strategy

However, the major obstacle to the success of this method is that it relied heavily on the fact that the tripeptide with toxin attached would have to be selectively coupled to the *N*-terminal of the CV-N(-3N). This is not a completely unfeasible requirement as *N*-terminal selectivity over lysine side chain amines is possible at a pH of 7 due to the difference in pKa's of lysine side chain amines and an *N*-terminal amine. Although initial peptide couplings were carried out to synthesise the terminal tripeptide, other strategies were developed that were deemed more achievable and so this strategy was discontinued.

2.3 Biolinker Development

A crucial aspect of attaching toxins to CV-N (**15**) is that once transported into infected cells through gp120 binding and recycling of the receptors there is a requirement for the toxin to be released in order to effectively kill the cell. Cell death may be possible with a non-biodegradable linker as the toxin may still be active even while attached to CV-N. Ideally though, the toxin should be released from the CV-N molecule once inside the infected cell.

Study in the field of polymer therapeutics has led to the discovery of several peptidyl linkers that are cleavable under low pH enzymatic conditions.[141] One such linker is the tetrapeptide Gly-Phe-Leu-Gly. Originally designed with treating cancerous tumours in mind, the linker will cleave inside the lysosomes of cells, which is at the low pH of 5.5, in the presence of pH-tolerant lysosomal peptidases such as cathepsin B.[141] The concept has been applied to this project, where once the CV-N-toxin conjugate is taken into the infected cell, cleavage of the biolinker will follow (**Figure 2.2**) releasing the toxin. This

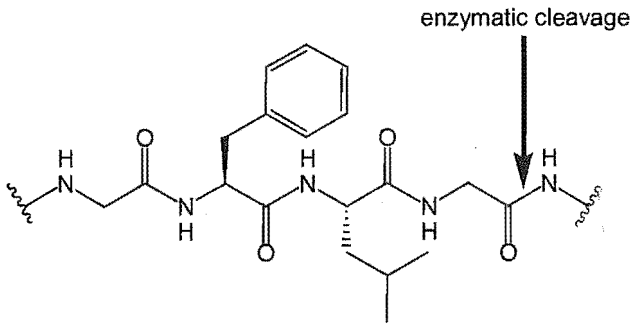


Figure 2.2 Biolinker

linker also acts as a spacer to separate the toxin from the surface of the CV-N molecule which is also crucial in allowing CV-N to bind gp120 effectively.

2.4 Succinimidyl Ester Strategy

Literature research suggested the most commonly utilised site of attachment of compounds to proteins was *via* the side chain of lysine residues present. The side chain amines of lysines are convenient nucleophilic functional groups that can be reacted with a variety of activated structures. Lysine is the preferred site of attachment as modification of other functional groups often leads to inactivation of the protein.[142]

The *N*-hydroxysuccinimidyl moieties were found to be such groups that are selectively targeted by primary amines with no interference with other functional groups.[143-147] Therefore, if these succinimidyl derivatives were to be incorporated into the overall structural design of the biolinker-toxin compounds, they would react with the *N*-terminal and/or the lysine side chains present in the CV-N (**15**) protein (**Figure 2.3**). This strategy means that multiple sites on CV-N (**15**) could be conjugated to toxin molecules. This has the advantage of ensuring more toxin is released into the cell, hence the opportunity for a quicker cell death. However, it also means that the products formed will not represent a homogeneous compound, but rather, a collection of CV-N molecules with varying numbers of toxin molecules attached.

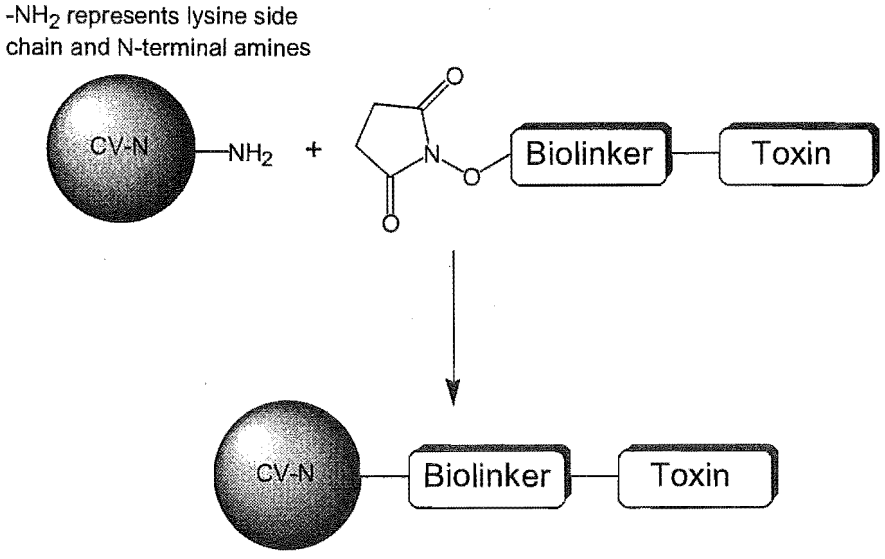


Figure 2.3 model for succinimidyl activation and reaction

2.5 Maleimides

2.5.1 CV-N-Cys Strategy

With the recombinant production of yet another CV-N variant at the LDDRD, this time with an extra cysteine residue present on the *C*-terminus, there arose the possibility of utilising the thiol side chain for attachment of conjugates. The native CV-N (**15**) already contains four cysteine residues but these are involved in internal disulfide bonds and unavailable for reaction. The presence of an additional cysteine residue provides the only free thiol present in the CV-N molecule. If conjugates could be linked through this thiol, the products would be homogeneous with the toxin attached at a defined position.

While looking at the feasibility of utilising succinimidyl esters, references to the use of maleimides through exclusive reactions with thiols was noted. However, without the presence of a free thiol group on the native CV-N (**15**) this method of attachment was not considered. Therefore, after the successful production of the CV-N-Cys construct it was

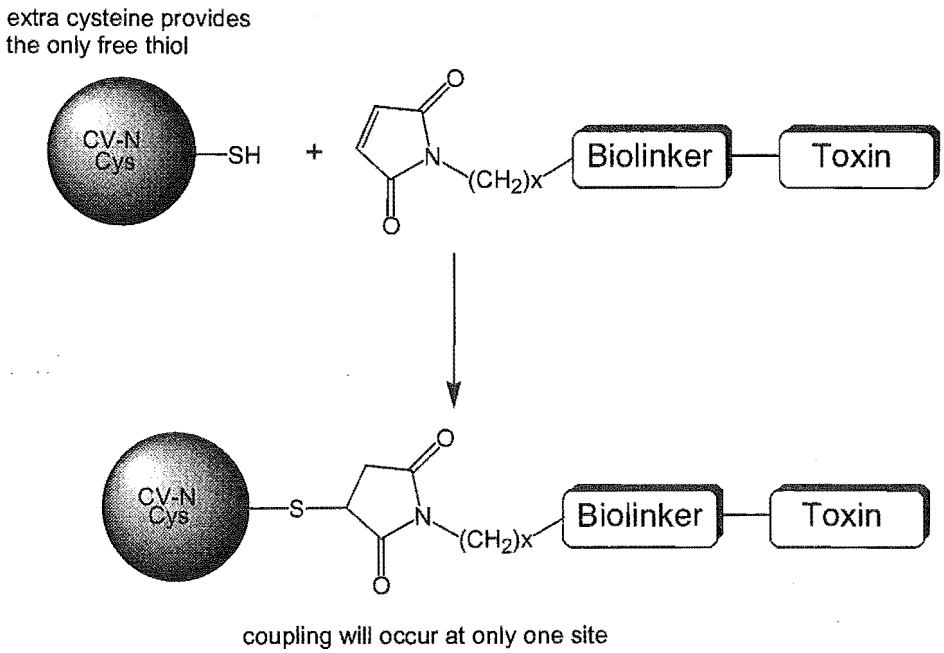


Figure 2.4 model for maleimide reaction with thiol

decided to develop methods to exploit the unique reactivity of the thiol towards maleimides.[148, 149] The reaction involves addition of the thiol across the double bond of the maleimide to yield a thioether (**Figure 2.4**). Maleimides do react with amines but usually require a higher pH than that for reaction with thiols.

2.5.2 Amine Conversion to Thiol Strategy

Recombinantly producing CV-N-Cys is not the only way to make available a free thiol on CV-N (**15**). There exist several reagents that selectively react with amines and convert them to thiols. With respect to CV-N (**15**) this converts the *N*-terminal and/or 5 lysine amines into sulfhydryl groups which become available to react with maleimides in the same way as CV-N-Cys above. A compound capable of this functionality conversion is 2-

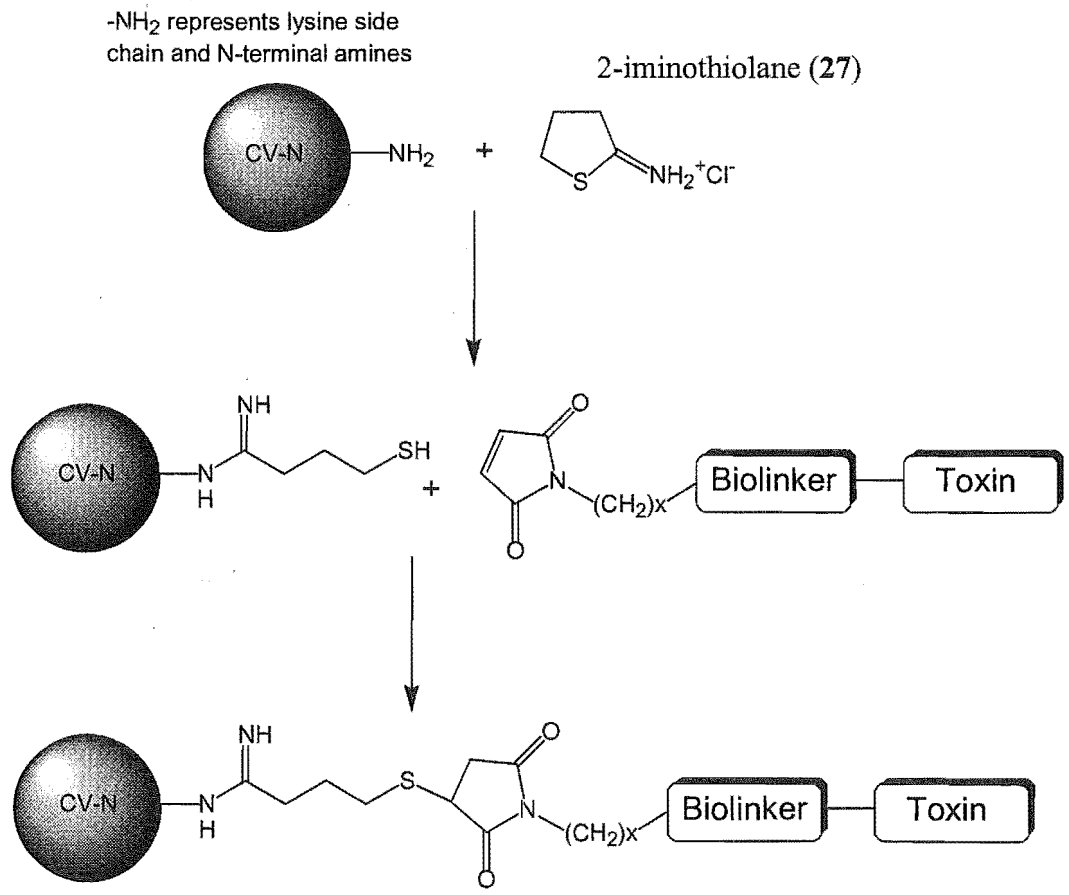


Figure 2.5 amine to thiol conversion strategy

iminothiolane (**27**).[150] This reaction opens the ring to form an amidine derivative with a pendant thiol ready for reaction with maleimides. The relative stability of 2-iminothiolane (**27**) in aqueous solution contrasts with the more rapid hydrolysis of other imido esters. This slow rate of hydrolysis makes it ideal for modification of proteins at low concentrations of reagent. **Figure 2.5** outlines the strategy developed to utilise this particular reagent in forming thiolated-CV-N and subsequent maleimide reaction. There are similarities with the CV-N-Cys approach with respect to the thiol-maleimide reaction, but in addition there is a distinct resemblance to the succinimidyl ester strategy in that there will be multiple sites of attachment for the biolinker-toxin conjugates.

2.6 Toxin Development

2.6.1 Mycalamides

Of the several toxin molecules available in the Marine Chemistry Group, two stood out as most suitable for use in this application. The first of these was the compound mycalamide A (**17**). The functionality available for conversion in this structure is the primary alcohol.

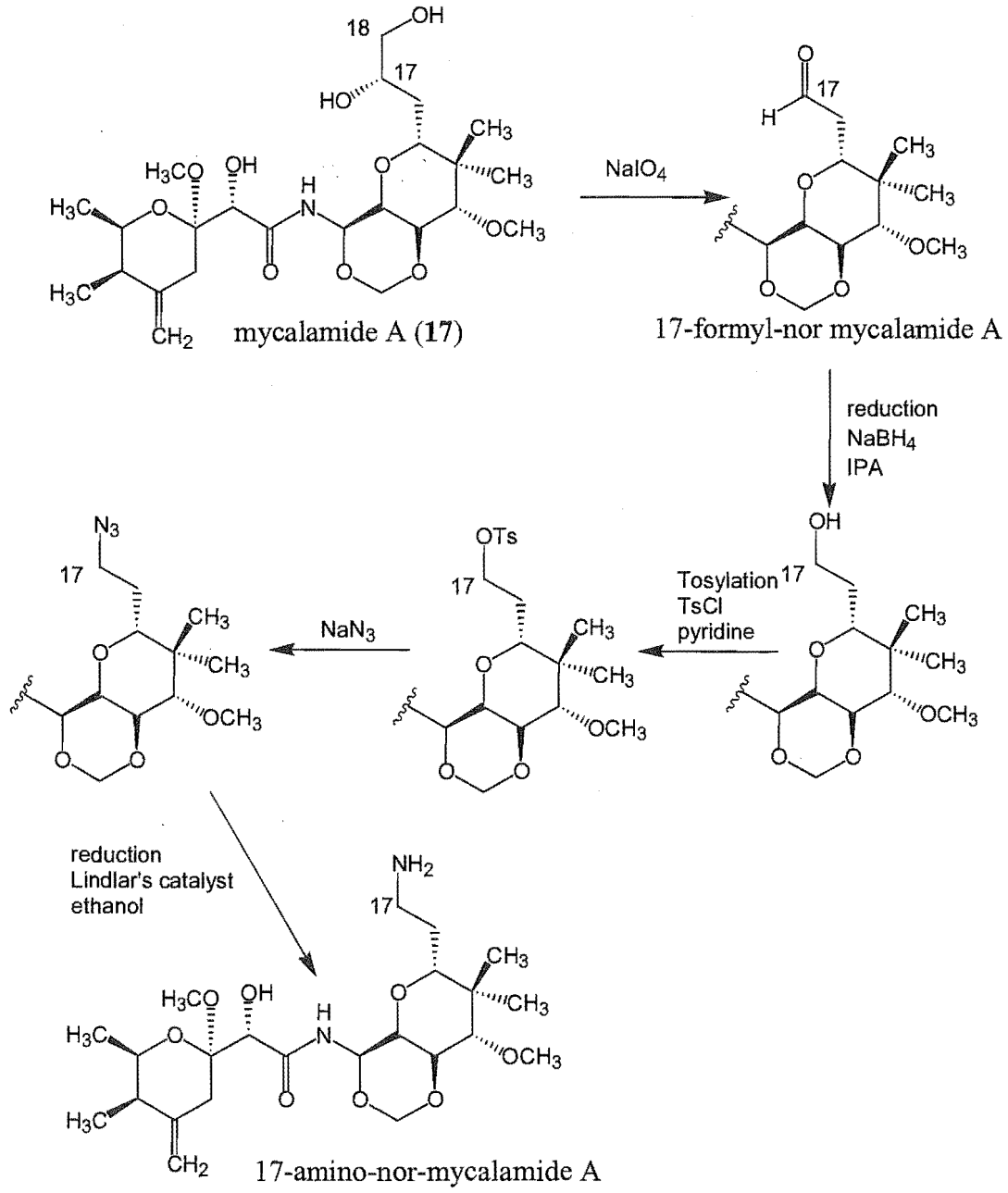


Figure 2.6 attempted mycalamide A (**17**) to amine conversion

Previous work done by Marie Squire in the Marine Group as part of an Honours project attempted to convert the alcohol to an amine in order to allow attachment to biolinkers.[151] The synthetic strategy is shown in **Figure 2.6**. The scheme shows initial oxidation of the diol (**17**) to an aldehyde and subsequent reduction to the single alcohol. Tosyl chloride converts the alcohol to the tosylate which allows transformation to the azide. Finally, reduction produces the amine. However, while this strategy worked up to the azide step, the final reduction yielded a product with a molecular mass two units higher than what was calculated for the amine. This result could not be rationalised at the time and so did not provide a suitable strategy to pursue.

An alternative, much simpler scheme was attempted by Sean Devenish as part of his Ph.D. research using reductive amination of the 17-formyl-nor-mycalamide A to give the desired amine.[152] While mass spectrometric analysis did show formation of the amine there was also a substantial amount of dimer present from the reductive amination of a 17-formyl-

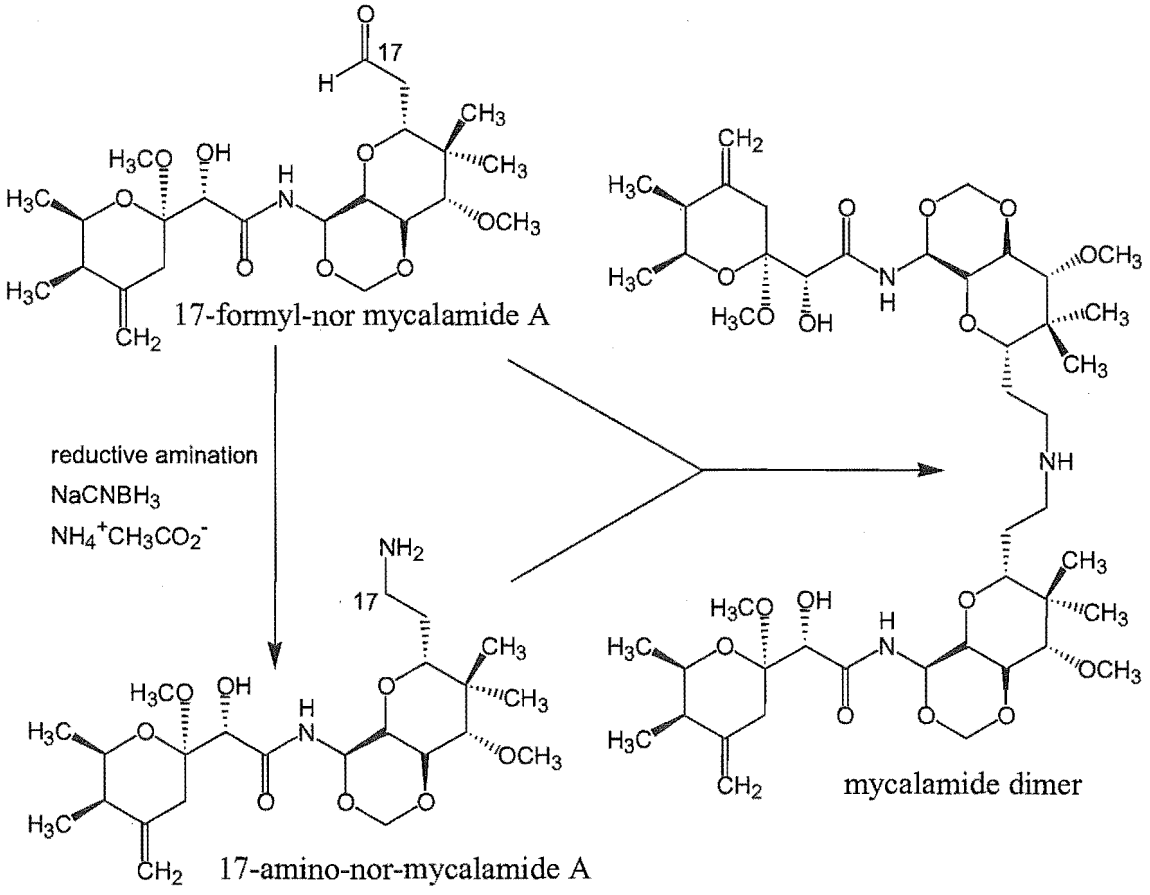


Figure 2.7 reductive amination of 17-formyl-nor-mycalamide A

nor-mycalamide A molecule with an already formed amine (**Figure 2.7**). This dimer did not possess a primary amine and as a dimer may not be as cytotoxic as the monomer. Various different reaction conditions and reagents were tested in order to prevent dimer formation, but with little success.[152] Work is continuing on alternatives to this reaction in order to be able to utilise the mycalamide class of toxins.

2.6.2 Halichondrins

A very similar set of reactions to those done on the mycalamide A toxin was performed on the halichondrin B molecule with more success. Dr Rachel Lill, as part of her Ph.D. thesis work, completed the series of reactions outlined in **Figure 2.8** leading to the synthesis of norhomohalichondrin B amine (**28**) from homohalichondrin B (**21**).[130] Although ultimately successful, this series of reactions was time consuming and difficulties in purification at each stage proved troublesome.

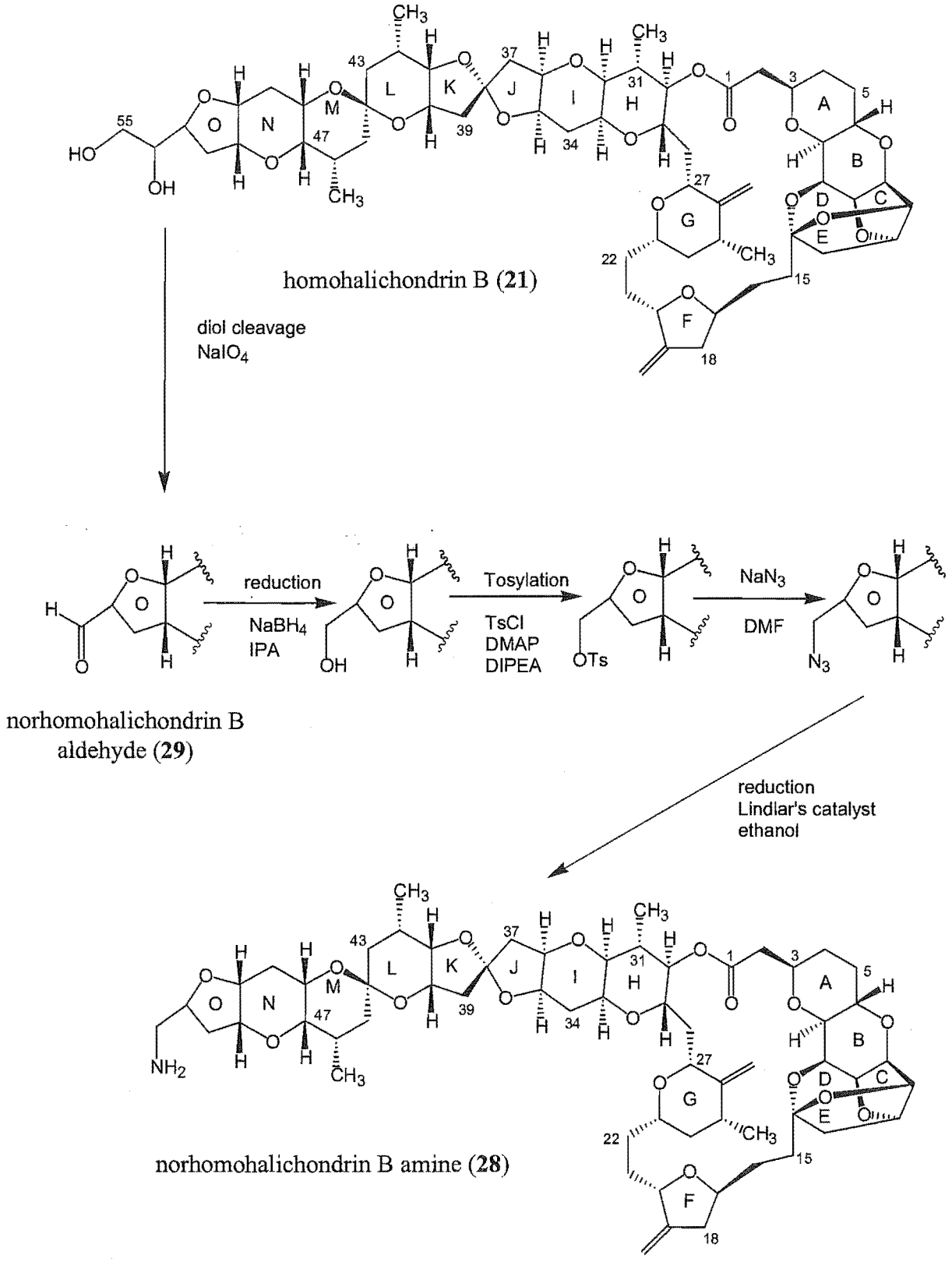


Figure 2.8 conversion of homohalichondrin B (**21**) to norhomohalichondrin B amine (**28**)

2.6.2.1 Reductive Amination of Norhomohalichondrin B Aldehyde

After the partial success of Sean Devenish's attempt at reducing the number of steps for this synthesis with respect to the mycalamide structure it was decided to attempt the same reductive amination on norhomohalichondrin B aldehyde (**29**) (**Figure 2.9**). It was hoped that the increased size of the halichondrin skeleton would prevent dimer formation. For

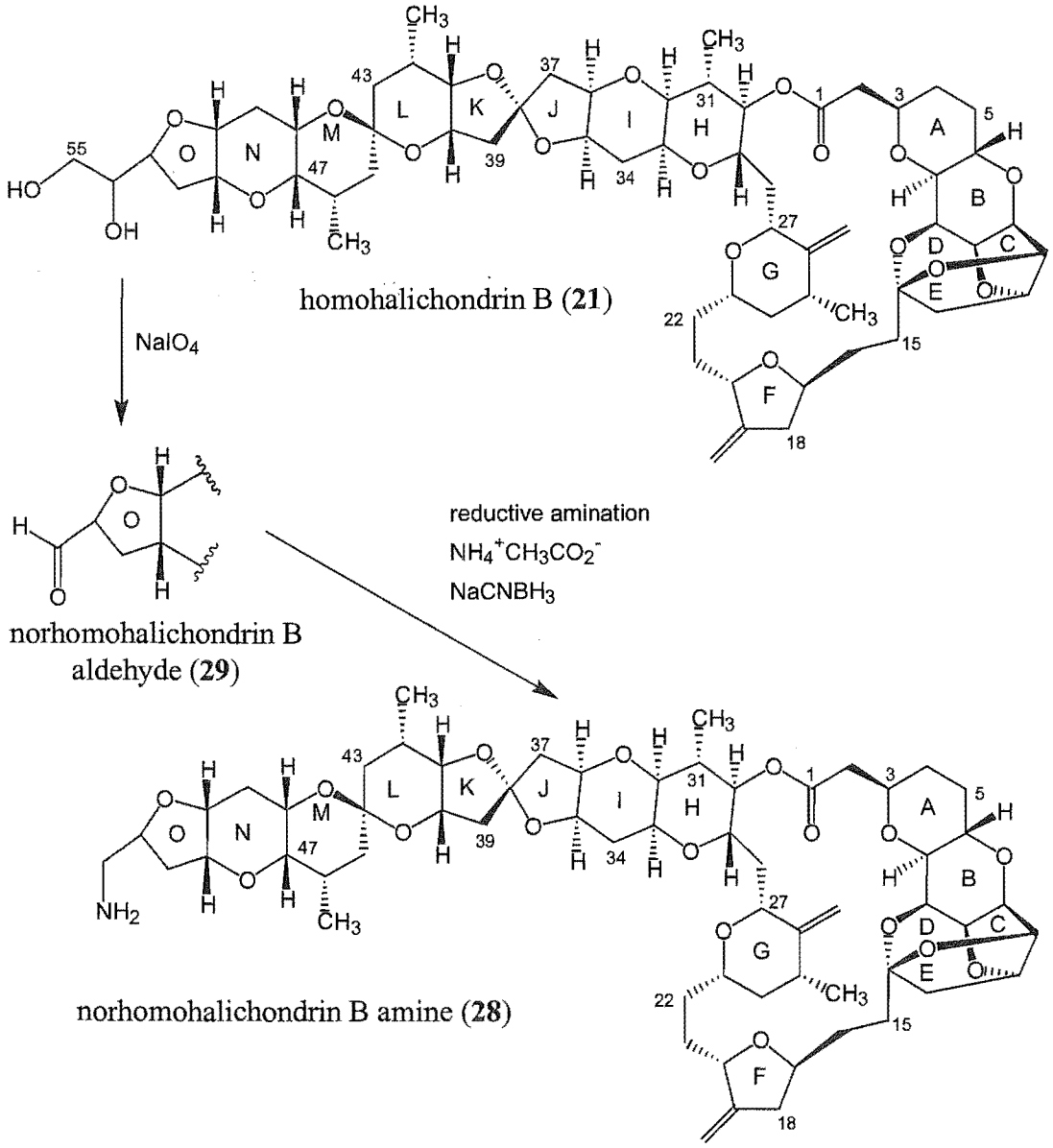


Figure 2.9 simple two step procedure to norhomohalichondrin B amine (**28**)

this to occur, two large halichondrin structures would have to meet in the correct orientation in order to react. This would be in competition with the reductive amination by the reagents sodium cyanoborohydride and ammonium acetate.

Norhomohalichondrin B aldehyde (**29**, 1.0 mg), ammonium acetate (100 equiv.) and $NaCNBH_3$ (0.7 equiv.) in MeOH were stirred together under argon at room temperature for 18 hours. Repeated TLC experiments demonstrated the formation of a much more polar product that was ninhydrin active. After complete disappearance of the aldehyde (TLC) the reaction mixture was purified using C18 column chromatography. A ^1H NMR spectrum was a very good match to that of Dr Lill's previous norhomohalichondrin B amine (**28**). Additionally there was no sign of the distinctive aldehyde resonance at δ 9.70 ppm. The most visible proton shifts of the 'O' ring of the aldehyde were the H53 and H50 proton resonances (shown in **Figure 2.10**) which had shifted from their initial values. Severe overlap in the 2D NMR spectra of the amine prevented the complete assignment of the 'O' ring by Dr Lill. More compelling evidence was obtained by ESIMS analysis with no sign of the MH^+ for the aldehyde (m/z 1091) nor the dimer (m/z 2199) as had been the case with mycalamide A (**17**).

This experiment was important as it circumvented the need for a five step synthetic scheme and allowed formation of norhomohalichondrin B amine (**28**) in two simple quantitative reactions with simple purification. As this reaction worked well and provided the amine efficiently, this toxin amine was chosen for subsequent attachment to biolinkers.

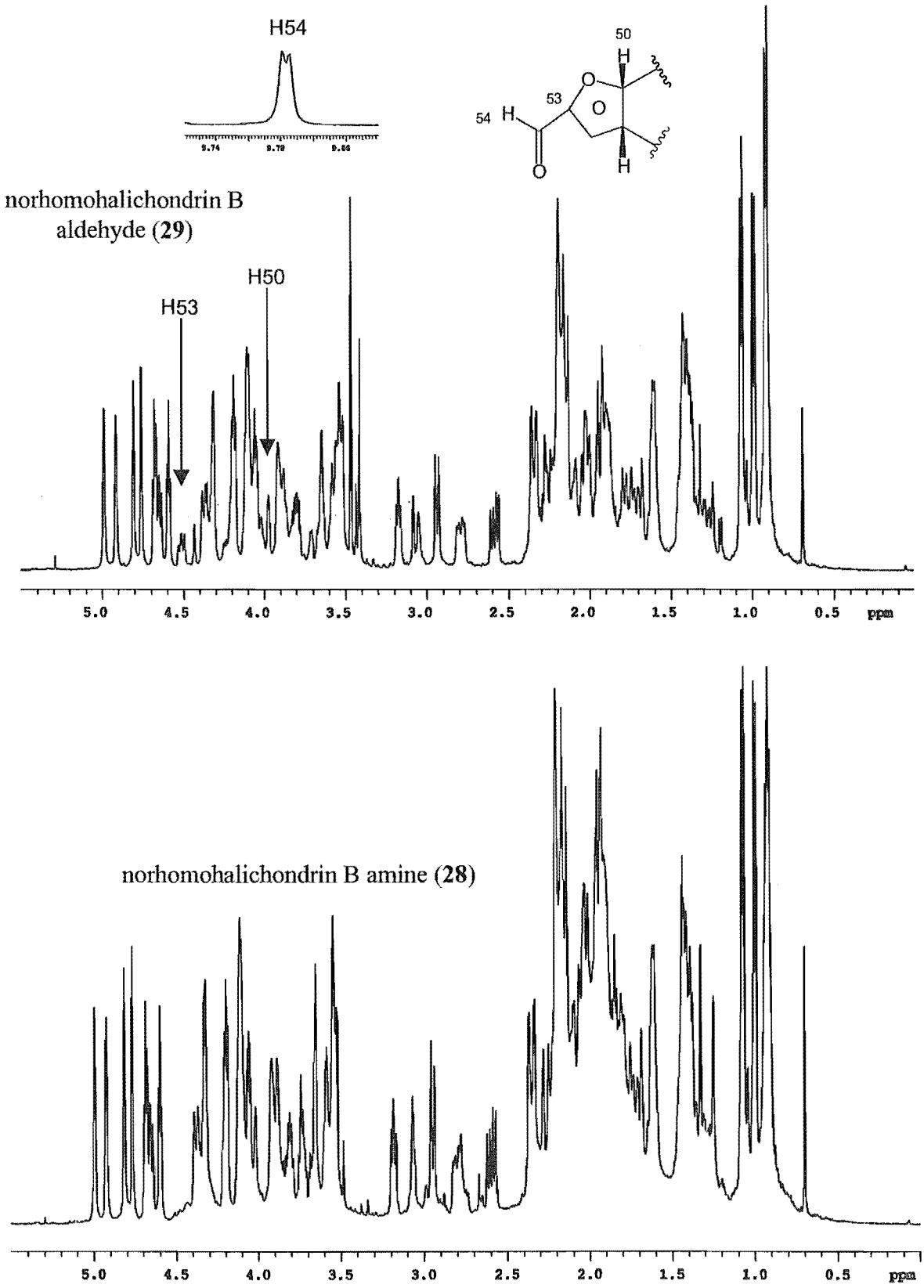


Figure 2.10 ^{1}H NMR comparison of norhomohalichondrin B aldehyde (**29**) and amine (**28**)

Chapter 3

PURIFICATION OF CV-N AND CV-N DERIVATIVES

3.1 Introduction

The delivery agent that provides targeting to HIV infected cells and virions is the CV-N (**15**) molecule. This whole project was based upon its efficacy in targeting gp120 expressed on the outer surface of infected cells. Therefore a supply of the native CV-N (**15**) and derivatives was essential in order to carry out the experimental aims of this thesis. There were ultimately three different CV-N type molecules utilised throughout this project. The first of these was the native CV-N (**15**) itself, then CV-N with an extra *C*-terminal cysteine residue (CV-N-Cys), and finally CV-N with lysine amines converted to thiols by iminothiolane (**27**) (thiolated-CV-N).

A crude batch of CV-N was kindly supplied by the LDDRD which represented side cuts from a previous large-scale purification of CV-N. A method of cleaning up the sample was needed to remove smaller proteins and other molecules from the mixture. Preparative HPLC was chosen as the best method to purify this crude CV-N.

The CV-N-Cys was produced recombinantly at the LDDRD while a guest in their laboratory. Utilising standard *E. coli* molecular biology techniques the CV-N-Cys protein was produced and purified. Mass spectrometry indicated the presence of a glutathione molecule attached to the free thiol rendering the product incapable of reacting with maleimides. Several methods of glutathione removal were carried out producing mixed

results. Subsequent experiments at the LDDRD using CV-N-Cys-glutathione did eventually successfully provide some CV-N-Cys.

The alternative method to produce thiols *via* amine conversion was at first carried out on a model protein, lysozyme. The concepts underlying these conversions were developed in co-operation with another Ph.D. student, Marie Squire. Analysis of this reaction of iminothiolane (**27**) with lysozyme allowed optimisation of the reaction conditions as well as optimisation of the mass spectrometry. Once satisfactory results were obtained with lysozyme, CV-N (**15**) was substituted in its place. Following size exclusion chromatography the products were analysed by mass spectrometry allowing an estimation of the degree of amine to thiol conversion. This thiolated-CV-N product was then ready for maleimide reaction.

3.2 Native CV-N Purification

A batch of CV-N side cuts was obtained from the LDDRD containing 1.712 g of semi-purified CV-N along with other smaller proteins and molecules. This mixture needed to be chromatographed in some way in order to separate out the pure CV-N (**15**) protein. The most obvious method of purification was by preparative HPLC. Successive analytical injections utilising various different gradient profiles demonstrated that a 20% to 50% CH_3CN/H_2O (0.05% TFA) mixture on a reverse phase C18 column gave the best separation and resolution of the CV-N (**15**) peak. Larger scale preparative HPLC runs were undertaken on a C4 column with the same solvent gradient profile as used analytically. Collections were made and the putative CV-N peak was reinjected analytically to confirm the success of the purification. The chromophore of CV-N (**15**) is shown in **Figure 3.1** and is typical of proteins in general as all residues contribute and produce an average UV spectrum. The slight maxima around 280 nm is due to the aromatic residues such as phenylalanine, tyrosine and tryptophan. A comparison of the crude, purified CV-N, and pure CV-N (supplied by the LDDRD) samples is shown in **Figure 3.2**. These HPLC traces demonstrate the effectiveness of this method for selectively collecting the compound of interest. This peak was also submitted for mass spectral analysis by LCMS which showed it represented a mass of 11,009 Da as expected for CV-N (**15**). A small shoulder peak of mass 6,584 Da was also noted but was estimated to represent <5% of the mass of the sample. Indications from the LDDRD had been that the crude CV-N was ~50% pure by mass, but after successive preparative HPLC collections the percentage purity was found to be only 12% by mass. A large proportion of

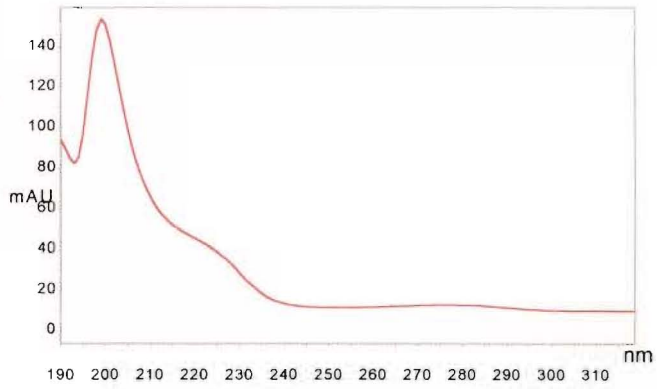


Figure 3.1 UV spectrum of CV-N

material was found to come off very early in the HPLC runs and therefore must have come from more polar compounds than CV-N (**15**).

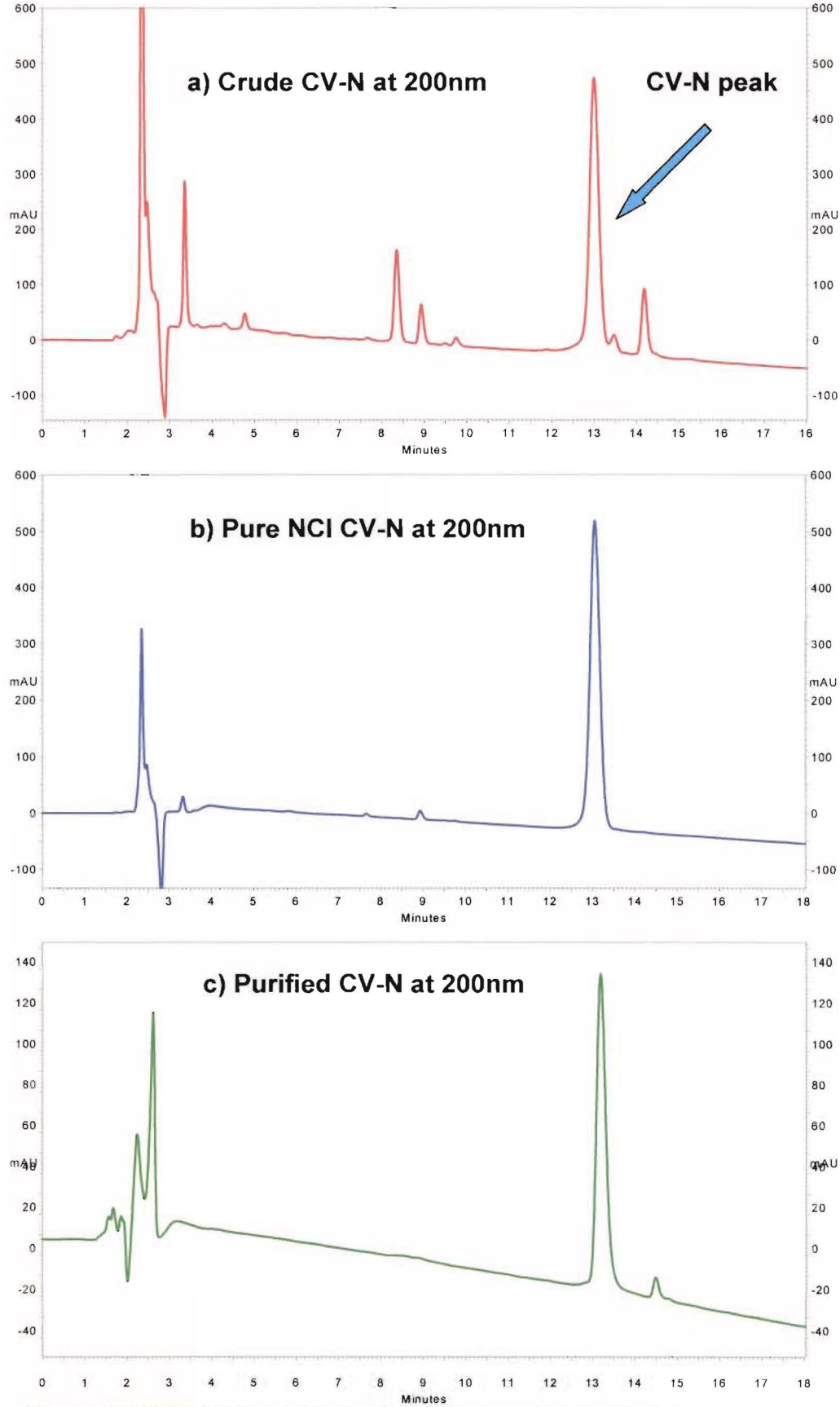


Figure 3.2 HPLC traces of **a)** crude, **b)** pure, and **c)** purified CV-N at 200 nm

3.3 CV-N-Cys

In comparing the strategies developed for this project, and discussed in chapter 2, the method employing the use of CV-N-Cys is probably the most desirable. This is because the product formed will be homogeneous, having only one toxin molecule per CV-N molecule. Therefore, production of CV-N-Cys was important in order for this strategy to be viable. Recombinant production of proteins from a DNA sequence is now standard practice around the world. The production of CV-N-Cys followed very similar protocols to those of native CV-N (**15**) reported by Boyd *et al.*[71] and Mori *et al.*[83] and described in section **1.2.3**. The basic idea, as outlined in **Figure 3.3**, is to introduce a plasmid (vector) containing the coding sequence for CV-N-Cys into *E. coli* cells through transformation. After incubation of the cells, the CV-N-Cys protein will be expressed through induction and then move into the periplasmic space. While cells exposed to slowly varying extracellular osmotic pressure are usually able to adapt to such changes, cells exposed to rapid changes in external osmolarity, can be mechanically injured. The procedure in this case was conducted by first allowing the cells to equilibrate internal and external osmotic pressure in a high sucrose medium, and then rapidly diluting away the sucrose. The resulting immediate overpressure of the cytosol is assumed to damage the cell membrane. Proteins released by this method are believed to be periplasmic, or at least located near the surface of the cell. In this way the CV-N-Cys is released into solution and chromatographic purification can take place.

3.3.1 Recombinant Production of CV-N-Cys

Whole *E. coli* cells containing the plasmid pET26b with coding for CV-N-Cys were added to a "Super Broth" solution made up of kanamycin (30 μg/mL), glucose (0.5%) and $MgSO_4$ (1.6 mM). This broth was incubated overnight at 37^{0}C. A sample of this inoculation broth was added to each of eight conical flasks (with fins for aeration) containing "Super Broth" (500 mL). These solutions were in turn incubated at 37^{0}C for four hours. An OD_{600}

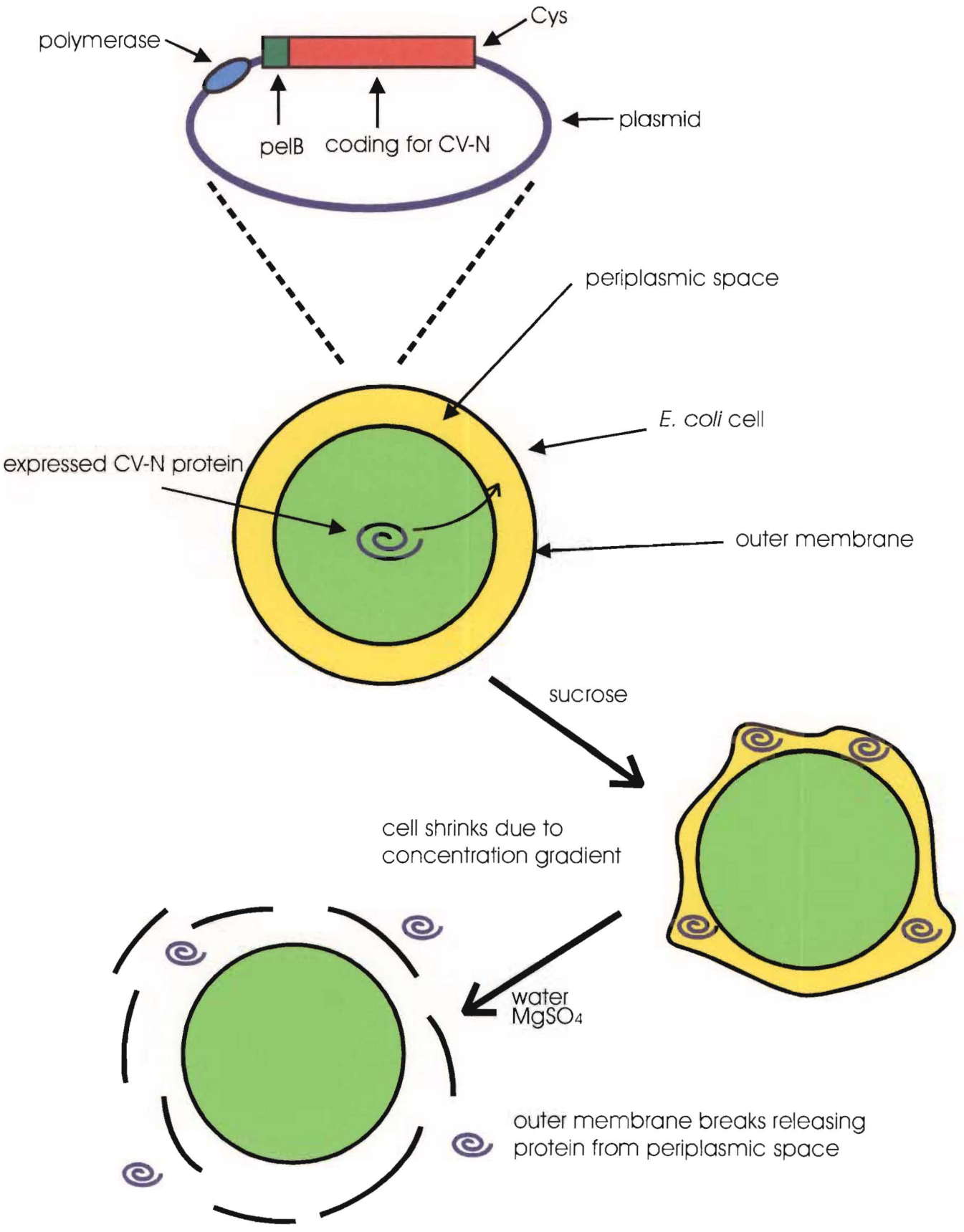


Figure 3.3 transformation and expression of CV-N-Cys protein in *E. coli* cells

(optical density) test indicated that the cells had grown sufficiently and were ready for induction. IPTG was added to each broth mixture (5 mL, 100 mM) in order to start the process of protein production. After another 2 hours incubation at 37^0C, the protein had been produced and needed purification. The combined broth solutions were centrifuged at 4^0C and the supernatants discarded. The remaining pellet containing the whole *E. coli* cells with CV-N-Cys incorporated in the periplasmic space was suspended in a high sucrose content solution and stirred for 15 min. Following centrifugation at 4^0C the supernatant was again discarded. The pellet was now suspended in ice cold $MgSO_4$ (5 mM) and again stirred for 15 min. This process burst open the outer membrane and released the CV-N-Cys into the surrounding solution. After centrifugation at 4^0C, the supernatant was retained this time as it contained the protein of interest. The solution was then frozen and freeze dried.

3.3.2 Purification and Analysis of CV-N-Cys

The freeze dried solid was resuspended in water and washed onto a reverse phase C4 pad. Washing with water removed salts such as $MgSO_4$ and other polar contaminants. An increasing gradient of methanol eluted the CV-N-Cys protein. Solvent was removed under vacuum. A Bradford Protein Assay (described in section **8.1**) estimated the mass of protein content to be approximately 75 mg at this stage. This value represented all proteins present, not specifically CV-N-Cys, and so would be reduced after further purification. At this stage, an SDS-PAGE (Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis) (described in section **8.1**) was also run in order to ascertain the relative molecular weights of all proteins present. SDS-PAGE utilises a voltage difference applied across a gel on which proteins migrate according to their size. Pre-determined size markers give an indication of the size of protein(s) present. In this case, the gel indicated the major component was a protein of approximately 11 kDa, but with many contaminants from proteins of various molecular masses with most being greater. This was expected at this stage of purification as the many proteins present in the periplasmic space would also be retained on the C4 pad and elute with the methanol-containing fractions.

The next stage was reverse phase C18 preparative HPLC with PDA detection at 210 nm. An increasing gradient from 20% to 50% CH_3CN/H_2O (0.05% TFA) eluted the major protein at approximately 35% CH_3CN. Collection of the peak over multiple injections yielded 43 mg of purified sample. A further SDS-PAGE experiment established that the sample had been purified with a single spot at approximate molecular weight 11 kDa. The correct mass for CV-N-Cys is 11,112 Da but an LCMS of the sample revealed a mass of 11,417 Da. This particular mass had been seen as a minor component with previous CV-N-Cys cultures at the LDDRD and determined to be CV-N-Cys-glutathione (**30**). Glutathione is a common cellular component containing a free thiol and during the expression or purification procedures, coupled to the free thiol presented by the extra cysteine residue to form a disulfide. However, in this case, all of the CV-N-Cys appeared to be coupled to glutathione. This was obviously unsuitable as the thiol side chain of the cysteine was to be utilised to react with maleimides.

3.3.3 Removal of Glutathione from CV-N-Cys-Glutathione

In order for the CV-N-Cys-glutathione (**30**) to be useful as a protein carrier the glutathione had to be removed. As the bond between the cysteine and the glutathione was a disulfide, the obvious method for breaking this bond would be reducing conditions. This would, however, also break the existing two disulfides present in the native CV-N structure. The early studies on CV-N (**15**) showed that a breakage of the internal disulfides led to a loss of anti-HIV activity.[71] This points to the disulfides having a crucial role in determining the active tertiary structure of CV-N (**15**). Therefore, under reducing conditions the anti-HIV activity would be lost. Ideally, a reagent that selectively cleaved a glutathione disulfide over a 'normal' cystine bond would be most useful, but such a reagent could not be found. Many reagents exist that do cleave all types of disulfide bonds such as dithiothreitol (DTT) and 2-mercaptoethanol. If these reagents were employed the protein would have to be allowed to refold back in the correct tertiary structure, forming the correct disulfides. Therefore, a strategy was developed to overcome the problem of glutathione addition by

cleaving all disulfides, removing the small molecules present, and finally allowing the CV-N-Cys to refold gently while reforming the native disulfide bridges.

3.3.3.1 Dialysis Treatment of CV-N-Cys-Glutathione

The first attempt undertaken for removal of the glutathione from CV-N-Cys utilised dialysis as the method for separation. The process was carried out using a 'Protein Refolding Kit' (Novagen) containing a dialysis cartridge, dialysis buffer and DTT (1 M). The cartridge contained a semi-permeable membrane with a MW cut-off of 2,000. Therefore, smaller molecules could transverse the membrane while larger molecules (such as CV-N-Cys) would be retained inside the cartridge. The driving force of this reaction was the change in concentration of the outer and inner solutions. The buffer in both inner (2.5 mL) and outer (1,000 mL) solutions was initially made up to 0.1 mM DTT therefore equilibrating both solutions. The DTT cleaved off the glutathione which then moved through the membrane to the larger volume. This would be due to the high concentration of cleaved glutathione inside the membrane which must equilibrate with the much larger outer solution, hence, effectively diluting it from the inner solution. This solution was left gently stirring at 4^0C for six hours and then the outer solution was changed to fresh dialysis buffer (0.1 mM DTT). After another 12 hours of stirring the outer solution was changed again, this time to just dialysis buffer. This would therefore permit any DTT remaining in the inner solution to permeate through the membrane and allow the CV-N-Cys, now free of glutathione, to refold. After another 6 hours the outer solution was refreshed with buffer and left stirring overnight. The inner solution was then removed from the cartridge and chromatographed on a small C4 column and washed with water to remove all salts and buffers present. Elution with methanol yielded the putative CV-N-Cys.

An SDS-PAGE revealed that when the sample was boiled prior to addition to the gel (as is standard for SDS-PAGE), the protein was a monomer of approximate molecular weight 11 kDa. If the sample was run without boiling, a dimer of weight 22 kDa was observed. An LCMS analysis also showed a molecular weight corresponding to dimeric CV-N-Cys.

Therefore the dialysis process had successfully removed the glutathione adduct but in the refolding process two monomers had come together to form a non-covalently bound dimer. Dimers of CV-N are not unusual as discussed in section **1.2.5** with respect to the solution and crystal structures of CV-N. The refolding conditions of the dialysis process must have been sufficient for the inducement of dimer formation. The dimeric CV-N-Cys was submitted for an XTT (anti-HIV) assay and showed an EC_{50} 12 times less than native CV-N (**15**) (submitted in tandem). However, at the level tested, a variation of only 12 times was not a significant difference. An EC_{50} is a measure of how well a drug will protect cells from infection by HIV virus. A thorough explanation of the XTT assay and the meaning of the EC_{50} result is discussed in section **6.2.2.** Together the analytical results show that the CV-N-Cys had more than likely refolded correctly (as the activity was still good) but in the process had formed a dimer.

3.3.3.2 Repeat of Dialysis of CV-N-Cys-Glutathione

The encouraging results above led to a second attempt to produce CV-N-Cys using dialysis to remove the glutathione. Several experimental changes were made, hopefully to improve the refolding section of the procedure and prevent formation of dimer. The initial dialysis treatment was the same as that above with the CV-N-Cys-glutathione placed in the dialysis cartridge with the inner and outer solutions set to 0.1 mM DTT. This was left stirring overnight. The inner solution was then transferred to 500 mL of water in order to dilute out the CV-N-Cys, allowing it to refold slowly. The solution was then progressively loaded into two (for balance) 3,000 MW cut-off Centriprep filters. These filters have a membrane which only allows molecules below a certain molecular weight cut-off to pass through the membrane, under centrifugal forces. As the initial solution passes through the membrane, more solution is added finally leaving a concentrated sample containing only compounds over MW 3,000, in this case CV-N-Cys. After this process the two samples were combined and freeze dried. The SDS-PAGE for this sample was similar to the first dialysis in that with pre-boiling of the sample it appeared as a monomer but without boiling the compound appeared as a dimer. The mass spectral analysis (LCMS) also revealed the

presence of dimer. The XTT assay result showed an EC_{50} 60 times greater than CV-N (**15**). Therefore, this experiment also successfully removed the glutathione molecule, but the protein also failed to refold as monomeric CV-N-Cys.

3.3.3.3 Chemical Removal of Glutathione

In an alternative approach to glutathione removal and protein refolding, CV-N-Cys-glutathione (10 mg) was dissolved in guanidine HCl (8 M) and vortexed gently before Tris HCl (33 μL of 3 M) and DTT (20 μL of 500 mM) were added to firstly denature the protein and secondly to cleave all disulfides present. Finally, the sample was flushed with N_2 gas, sealed with foil and left in the dark for 2 hours. A G100 column (40 x 2.5 cm) was prepared eluting with PBS (Phosphate Buffered Saline). The sample was chromatographed and fractions collected every 2 min with a flow rate of 2.1 mL/min. Monitoring at 280 nm allowed collection of a single peak over 10 fractions from 32-52 min. The combined fractions were filtered through a 3,000 MW Centriprep as for the previous sample above and concentrated to a small volume (~10 mL) and then freeze dried. SDS-PAGE showed monomer with boiling of the sample, and approximately a ratio of 1:2 monomer to dimer without boiling. In contrast to the dialysis experiments the LCMS of this sample showed distinctly monomer with a mass of 11,112 Da. However, the XTT assay showed an EC_{50} 668 times greater than CV-N (**15**) which means that this sample is significantly less active. Therefore, this method also removed the glutathione successfully and importantly produced CV-N-Cys monomer (LCMS), but on refolding had obviously failed to form the correct tertiary structure.

3.4 CV-N Amines Converted to Thiols

3.4.1 Lysozyme Model with Iminothiolane

In order to test the efficiency of converting a protein's lysine side chain amines to thiols with iminothiolane (**27**), a model protein (lysozyme) was first tested. This model system would form the basis of subsequent reactions with CV-N (**15**) to perform the same functional group transformation as outlined in section **2.5.2**. Lysozyme has a molecular weight of approximately 14,300 Da and so was similar in size to CV-N (**15**) therefore making it a suitable model. More importantly lysozyme has eight lysines and an *N*-terminus therefore giving nine potential sites for conversion as compared to CV-N's (**15**) six amines.

Lysozyme (1 mg) was dissolved in reaction buffer (1 mL, 0.1 M sodium borate, 0.001 M EDTA, 0.15 M NaCl) and 10 equivalents of iminothiolane (**27**) added (96 μL of 1 mg/mL solution). The reaction was followed by ESIMS analysis with 5 μL aliquots diluted 30-fold and injected directly into the mass spectrometer with 5 μL of a 10% formic acid solution to protonate the protein. The spectrum from the first injection, 5 min after the iminothiolane (**27**) was added, is shown in **Figure 3.4a**) and is typical of proteins.

In order to show a signal in an electrospray mass spectrometer the compound of interest must have the ability to be protonated for positive mode electrospray, or lose a proton for negative mode work. Proteins have many functional groups that can be protonated and therefore can be readily multiply charged. The signal from the mass spectrometer is plotted as an intensity *versus* mass/charge (m/z) ratio. So, for example, if a compound had a molecular mass of 10,000 Da and was capable of holding 10 charges (10 H^+) a signal would be detected at 1,001 mass units. Therefore, in this case the 'A' series ions (blue) in **Figure 3.4a)**, representing native lysozyme, are shown with charges ranging from 9 (m/z = 1590.3) to 17 (m/z = 842.4). The 'B' series (red) showed the beginnings of formation of

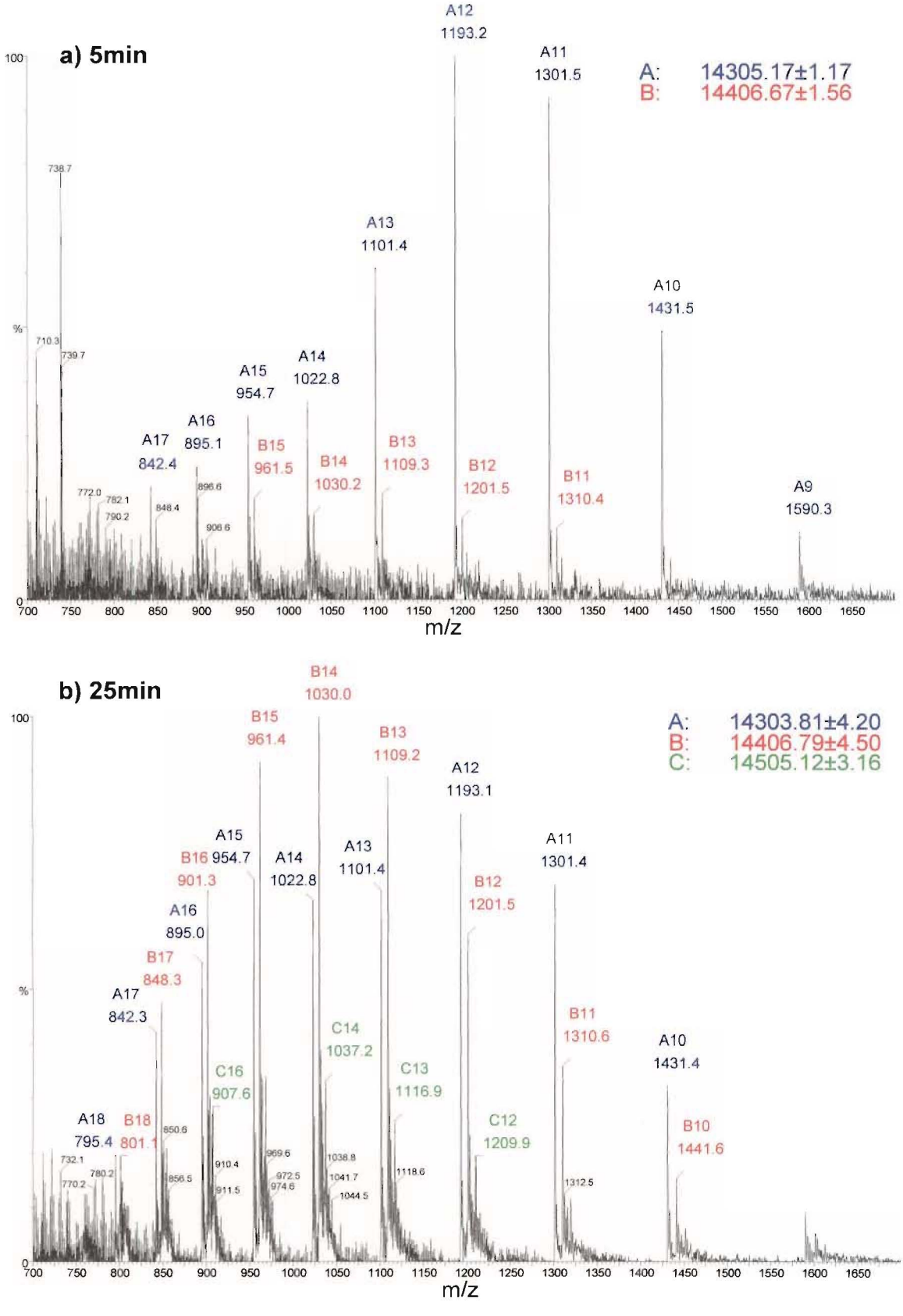
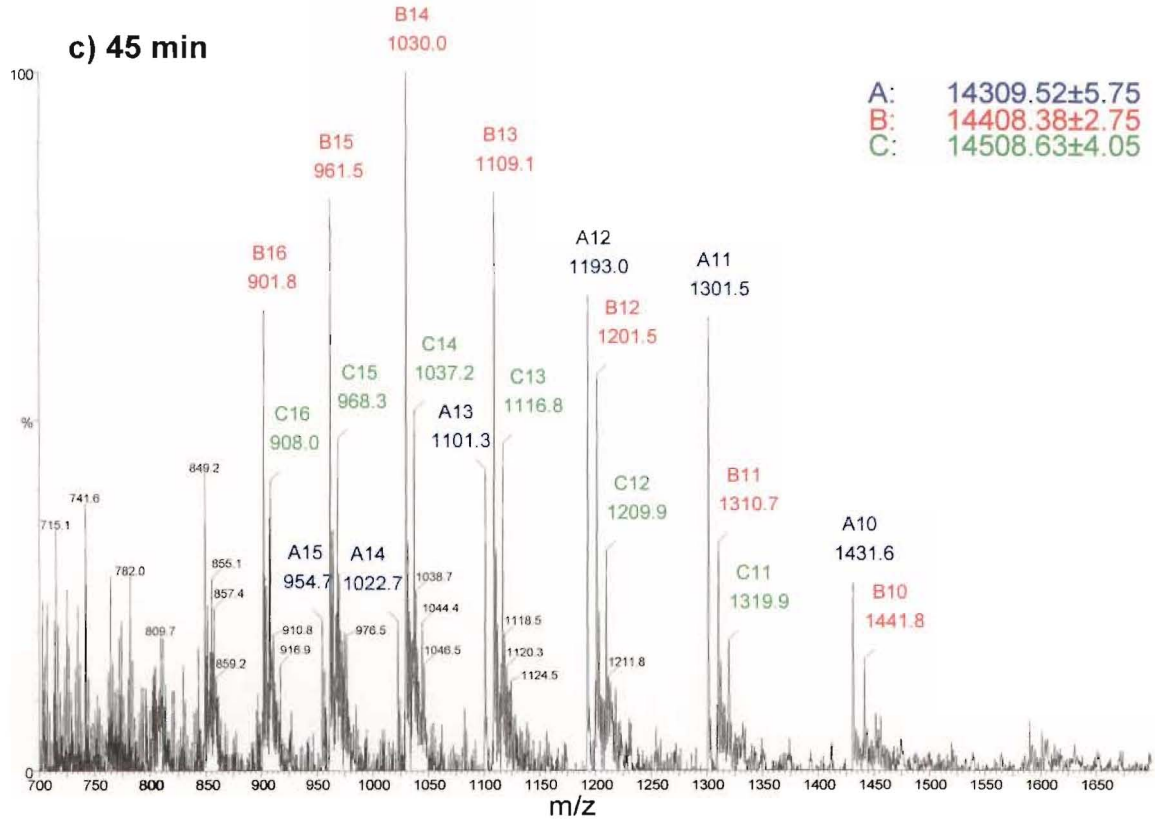


Figure 3.4a)-b) ESIMS spectra of lysosyme plus iminothiolane (**27**) after **a)** 5min & **b)** 25min



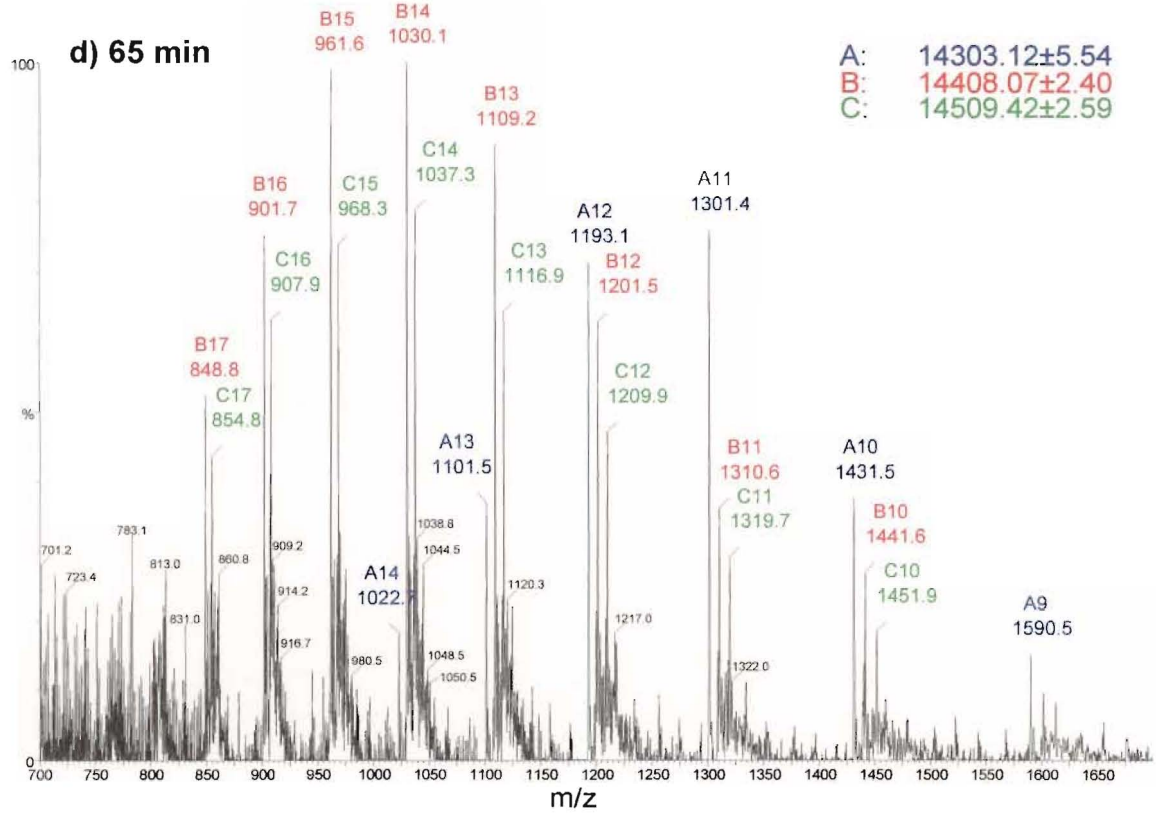


Figure 3.4c)-d) ESIMS spectra of lysosyme plus iminothiolane (**27**) after **c)** 45min & **d)** 65min

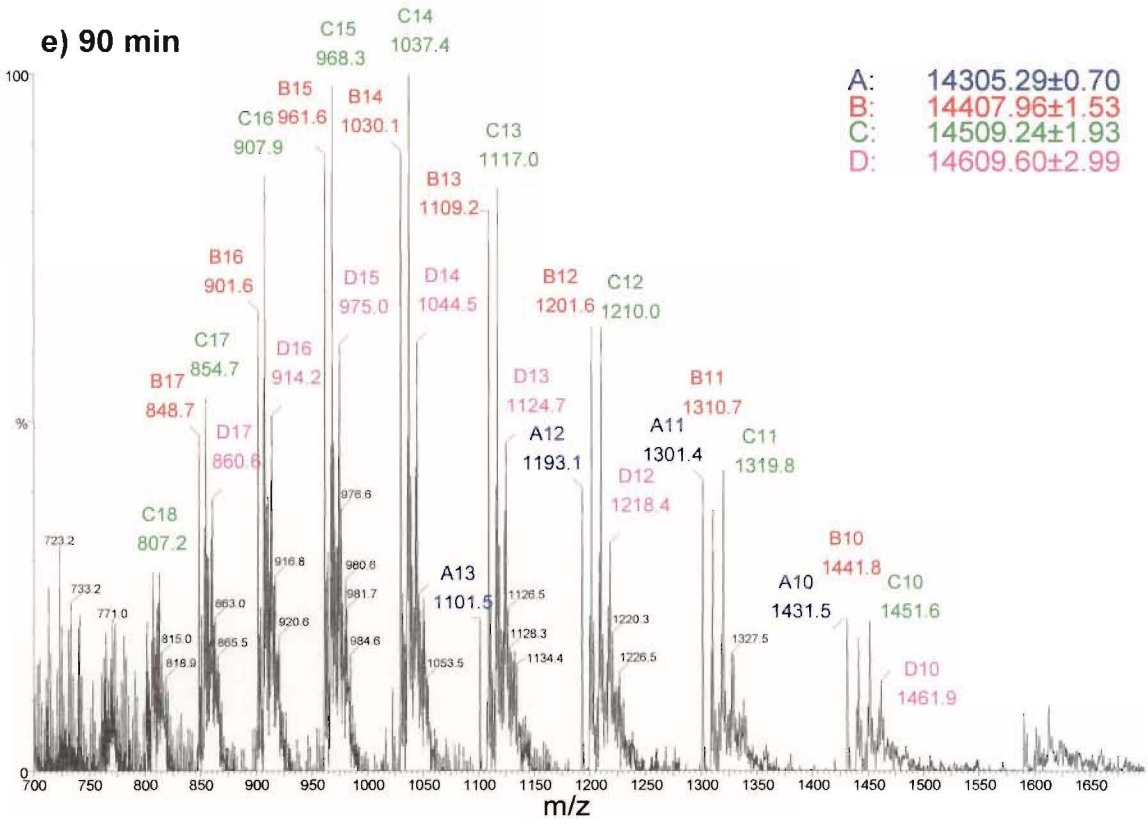


Figure 3.4e) ESIMS spectrum of lysosyme plus iminothiolane (**27**) after **e)** 90min

lysozyme protein with one single iminothiolane molecule attached. The numbers in the top right hand corner of **Figure 3.4a)** represent an estimate of the total molecular mass of the protein [calculated by multiplying the particular ions by their charges and then taking away the number of charges (as a charge is due to a proton of mass ~1)]. The numbers are averaged and the error is shown to the right of the mass. As the peaks become smaller or more concentrated they are less easily resolved and the error in the calculation increases to reflect this.

The addition of a single iminothiolane molecule to form the amidine derivative of lysozyme increases the mass of the protein by 101 mass units and therefore raises the overall mass from ~14,305 to ~14,406 Da. The next injection, 25 min after addition of iminothiolane (**27**) is shown in **Figure 3.4 b)**. Now the 'B' series of ions has surpassed the 'A' series in intensity confirming that lysozyme has had a single iminothiolane molecule

added. There was also the beginnings of a series of ions ('C', green) that represented addition of two iminothiolane molecules to lysozyme. After 45 min, (**Figure 3.4c**)) the dominant species was lysozyme plus one iminothiolane ('B' series) while the plus two ('C' series) had almost reached the level for that of the native lysozyme ('A' series). The next spectra at 65 min (**Figure 3.4d**)) continued the same trend with the 'B' series dominant, then 'C' and 'A'. The final spectrum at 90 min (**Figure 3.4e**)) showed the lysozyme plus 2 iminothiolane molecule(s) as the most abundant species while the remaining native lysozyme had dropped below the observed level of the +2, +3, and +4 additions.

This experiment confirmed that the reaction of iminothiolane (**27**) with a protein converted the lysine side chain amines to thiols in a time-dependent manner. It also demonstrated how readily the reaction can be followed by mass spectrometry. Importantly, there is minimal sample wasted by analysis.

The reaction mixture was then further reacted with maleimide derivatives to test the next stage of the reaction scheme. This is detailed in section **5.2.1**.

3.4.2 CV-N Plus Iminothiolane

3.4.2.1 Preliminary Reaction to Test Effectiveness of Thiolation

The positive result with the model protein lysozyme led to the same reaction being performed on CV-N (**15**). The first attempt at this reaction used 200 μg of CV-N (**15**) with 10 equiv. of iminothiolane (**27**) but yielded different results from that observed for lysozyme. The reaction mixture was analysed by taking 10 μL aliquots, diluting them followed by ESIMS analysis. This was carried out after 5 min and 25 min. The 5 min sample merely showed native CV-N (**15**) with no modified protein visible in the mass spectrum. The 25 min sample showed only a very small conversion of CV-N to the plus 1 iminothiolane product. A further 20 equiv. of iminothiolane (**27**) was added to force the

reaction to proceed. After another 5 min reaction period, ESIMS analysis showed formation of the single iminothiolane addition product. At a total reaction time of 45 min the analysis indicated plus 1 and plus 2 iminothiolane adducts present. Addition of a third iminothiolane was apparent after 60 min although the reaction rate was much slower than for lysozyme. There was still a large proportion of native CV-N (**15**) present, therefore another 20 equiv. of iminothiolane (**27**) was added to the reaction mixture. The 80 min spectrum (**Figure 3.5**) showed approximately equal proportions of native CV-N (**15**) to the plus 1 iminothiolane adduct with smaller proportions of the plus 2 and plus 3. At this stage the mixture was reacted with a maleimide group (see section **5.2.2**).

In comparison to lysozyme there were distinct differences to this situation. All reaction conditions remained identical, but the CV-N (**15**) protein was much slower to react with the iminothiolane (**27**). However, the main objective was achieved as reaction did take place, with CV-N (**15**) requiring a greater number of equivalents of iminothiolane (**27**) and/or a longer reaction time compared to lysozyme.

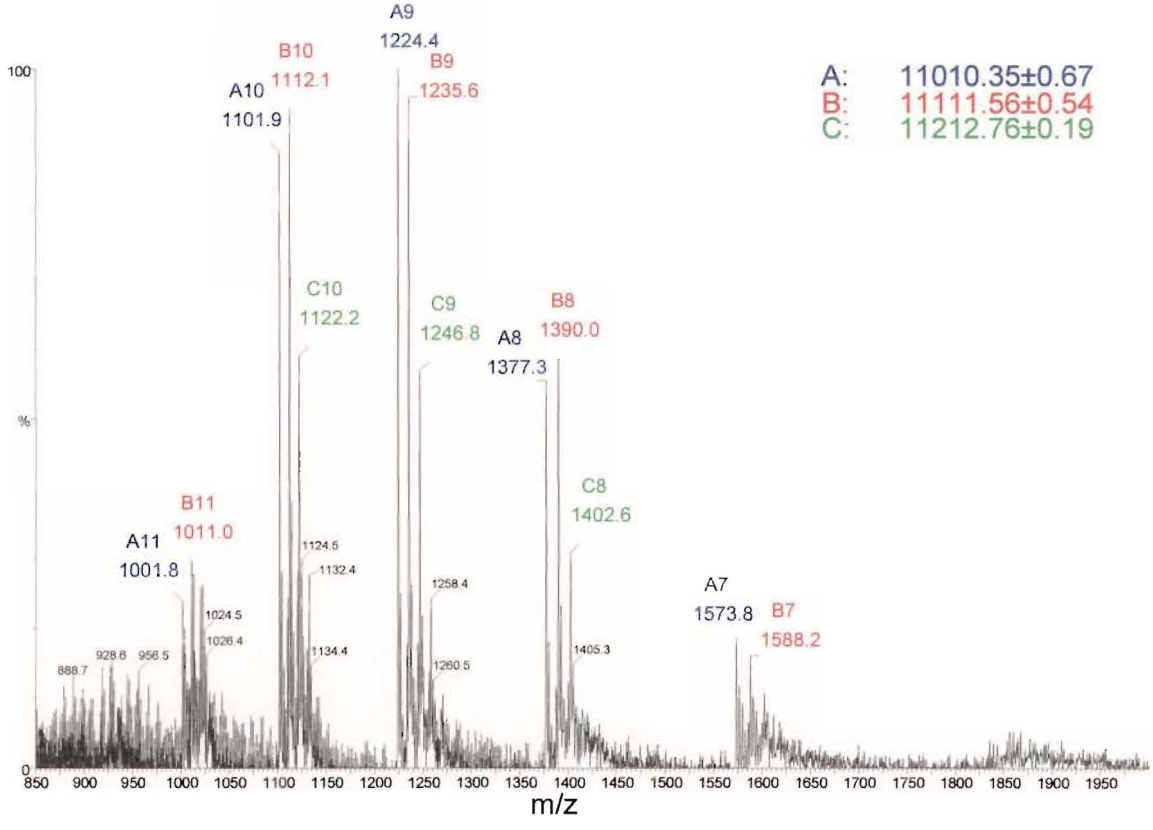


Figure 3.5 ESIMS spectra of CV-N plus iminothiolane (**27**) after 80min

3.4.2.2 Second Reaction with Modified Conditions

The reaction between CV-N (**15**) and iminothiolane (**27**) was re-examined over a longer time frame and using a higher ratio of iminothiolane (**27**) to CV-N (**15**). CV-N (**15**, 200 μg) was taken up in reaction buffer and 50 equiv. of iminothiolane (**27**) added. Monitoring by ESIMS analysis at intervals did indeed show formation of the desired iminothiolane adducts. The reaction was still somewhat slower than that for lysozyme, but the formation of significant amounts of thiolated-CV-N was observed by ESIMS analysis.

With the large excess of iminothiolane (**27**) present it became necessary to purify the CV-N adducts prior to reaction with maleimides. Although iminothiolane (**27**) reacts preferentially with amines it is also slowly hydrolysed to a free thiol. With a slow reaction and a high ratio of iminothiolane (**27**) to protein there is a significant concentration of free thiol originating from the iminothiolane. This would interfere with reaction of any maleimide derivative with the protein thiols necessitating a purification step. The thiolated-CV-N adducts were chromatographed on a G25 size exclusion column eluting with water with monitoring at 276 nm. The protein adducts eluted first followed by the lower molecular weight components present including excess iminothiolane (**27**) and its hydrolysed form. After concentration of the protein fraction, the thiolated-CV-N was ready for derivatisation with maleimide containing constructs (section **5.3.2**).

3.4.2.3 Optimised Reaction of CV-N Plus Iminothiolane

This final reaction between CV-N (**15**) and iminothiolane (**27**) took into account the previous reactions in order to optimise the attachment of the thiol group to the protein. To CV-N (**15**, 200 μg) in reaction buffer (0.1M sodium borate, 0.001M EDTA, 0.15M NaCl) was added 100 equiv. of iminothiolane (**27**). The start of formation of the thiolated-CV-N was noted after ESIMS injections at 5 min and 25 min respectively. The signal to noise ratio was quite poor mainly due to the presence of the buffer suppressing the ionisation of

the protein. However, the product formed was primarily of the +1 with a little +2 adduct and therefore the reaction was stopped. This was achieved by adding the reaction solution to a G25 chromatographic column eluting with water 30 min after the reaction had begun. The column separated the protein material from the small molecule buffering ions and excess iminothiolane (**27**) present. The protein fraction was collected and with the removal of the small molecules, easily analysed by ESIMS (**Figure 3.6**). The reason the reaction was stopped and purified at this stage was to provide primarily one site on the CV-N protein for modification with maleimides. The 100 equiv. of iminothiolane (**27**) (as opposed to the previous 50 equiv.) allowed the reaction to proceed faster with the G25 chromatography removing the excess reagents effectively. If a larger degree of modification was required then the reaction would merely need to be left for longer with monitoring by ESIMS injections at various time intervals. The thiols formed were relatively stable in that they survived the column chromatography and it was noted that they can be kept in solution for at least 24 hours although ideally they should be formed and reacted as needed. This particular thiolated-CV-N was reacted with a maleimido toxin derivative in section **5.5.2**.

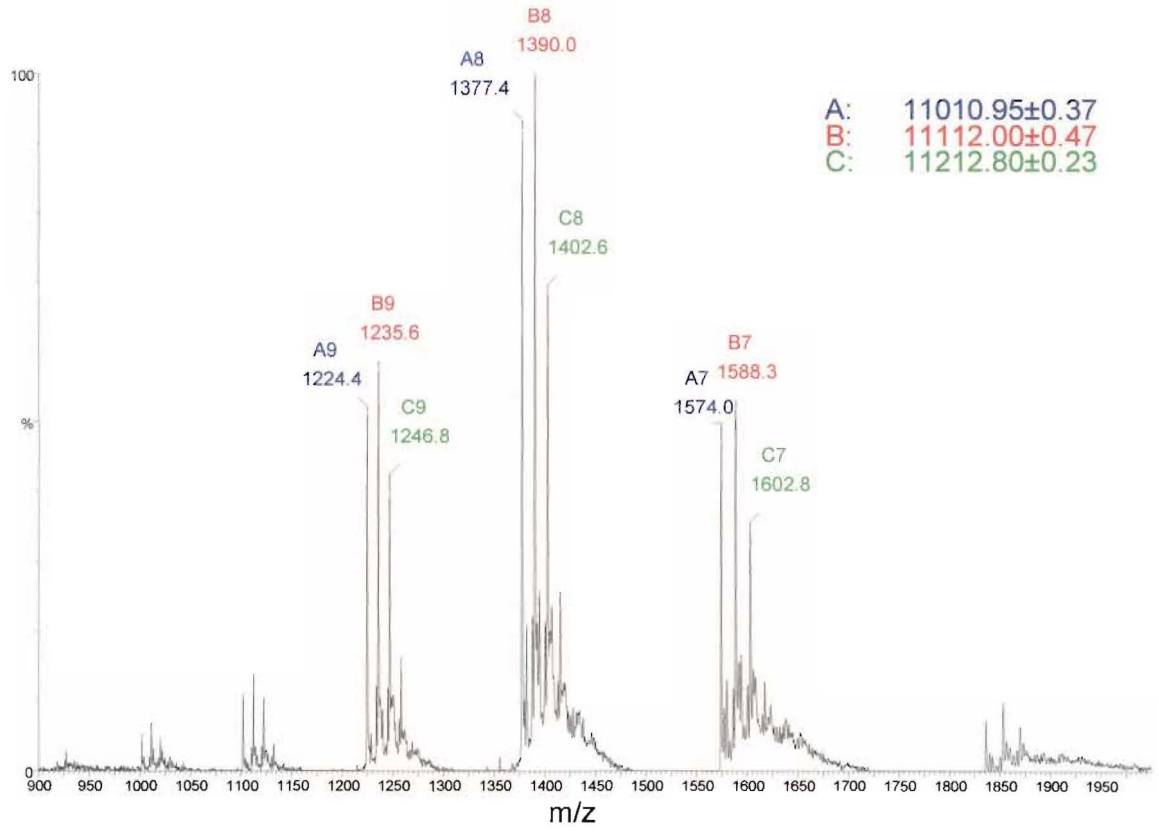


Figure 3.6 ESIMS of CV-N plus iminothiolane (**27**)

3.5 Conclusions from CV-N Development

There were three CV-N derivatives developed as part of this project. The first was the native CV-N (**15**) itself. Although a lengthy and time consuming process, preparative HPLC did purify the CV-N (**15**) protein to an acceptable level. The second derivative, CV-N-Cys, proved much more difficult to obtain as, although the recombinant production proceeded smoothly, the final product had glutathione attached. Removal of this impurity was achieved although refolding of the protein caused either dimer formation or loss of bioactivity. The last derivative utilised in this project was thiolated-CV-N obtained through reaction of CV-N (**15**) with iminothiolane (**27**). After initial testing with lysozyme, the transformation was carried out on CV-N (**15**) to produce the protein with the presence of free thiols.

If the CV-N-Cys could be produced free of impurities then its manufacture and purification would be identical to that of the native CV-N (**15**) originally produced. The thiolation was a single quick step from the native protein and so overall there would be very little difference in long term supply or production of these three CV-N derivatives. The differences lie in the fact that the native protein can have conjugation in up to six sites, the CV-N-Cys at a single position while the thiolation can produce whatever degree of conjugation is required between these numbers. Therefore, although very similar in structure, these three CV-N derivatives have distinct profiles that can be exploited to produce specific target conjugates.

Chapter 4

SUCCINIMIDYL METHODOLOGY

4.1 Introduction

The succinimidyl entity is the most commonly utilised grouping for attaching small molecules to proteins. The rationale for this usage is the relative stability of the succinimidyl group in solution and selective reactivity towards primary amines. The introduction of the succinimidyl group can be through either a carboxyl, amine or alcohol functionality on the compound to be activated, which is again a demonstration of the utility of this activation method. This chapter describes three areas where succinimidyl ester methodology was utilised.

Many of the derivatising reagents available for protein modification exploit the reactivity of the succinimidyl ester. Two such compounds are the fluorescent reagents 5-SFX (**31**), and BODIPY FLX-SE (**32**), both of which were chosen to react with CV-N (**15**) to test a variety of reaction parameters before commitment of toxin derivatives. These activated dye molecules did not contain a cleavable biolinker and as such did not specifically test that phase of the reaction scheme. However, their reaction profiles and subsequent analysis allowed formulation of a strategy when dealing with the subsequent biolinker derived conjugates.

Before reaction of CV-N (**15**) with toxin containing conjugates, the synthetic scheme leading to these final products needed to be tested to optimise each reaction step. The choice was made to substitute the toxin for a fluorescent derivative. This derivative would have the advantage of being chromatographically distinct by UV profile and ultimately lead to fluorescent conjugates of CV-N (**15**) that could be more easily detected and tested. The coumarin molecule is a well-characterised fluorescent compound and was readily available with a free primary amine as is the case with toxins such as norhomohalichondrin B amine

(**28**). A final strategy was developed that would be suitable for replacement of the coumarin with the halichondrin moiety.

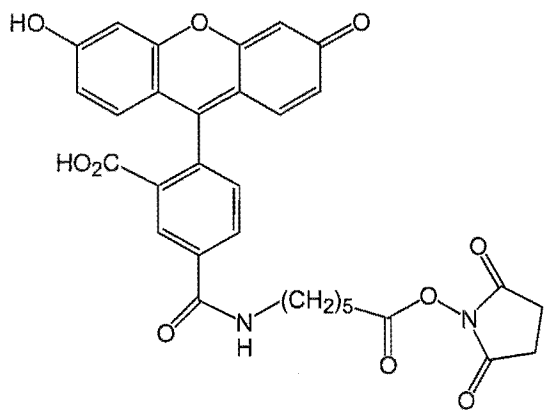
The last section of this chapter deals with the synthesis of succinimidyl ester biolinker-toxin conjugates in order to activate them for reaction with CV-N (**15**). The synthetic strategy was again developed from the reactions involving the coumarin conjugate in overcoming various problems. The final synthesis was successful and yielded a CV-N conjugate with a biolinker and toxin attached.

4.2 Fluorescein 5-SFX and BODIPY FLX-SE

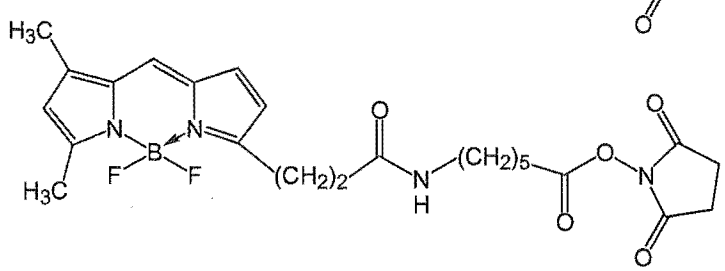
4.2.1 Dye Reactivity

The structures of the two dye molecules (Molecular Probes Inc.) are shown below. Both dyes have a $(CH_2)_5$ linker separating the 'fluorescent' section of the molecule from the succinimidyl ester. These were specifically chosen to mimic the relative size of the peptide biolinker that would ultimately be present in the toxin conjugates. This pendant chain also separates the fluorophore from its point of attachment, potentially reducing the interaction of the dye molecule with the protein after conjugation and minimising interference of the binding of the CV-N derivative to gp120. The molecular weights of the two dyes are 587 Da for 5-SFX (**31**) and 502 Da for the BODIPY conjugate (**32**). Once attached to the side chain amines of CV-N (**15**), the increase in mass per addition would be 471 Da and 387 Da respectively, allowing easy identification of the CV-N-dye conjugates.

Fluorescein 5-SFX (**31**)



BODIPY FLX-SE (**32**)



4.2.2 5-SFX Reactions with CV-N

4.2.2.1 Reactions at pH 8.2

The reactions of 5-SFX with CV-N (**15**, 200 μg) was carried out on in $NaHCO_3$ buffer (200 μL, 0.1 M) at room temperature. The pH of the buffer was 8.2 which is ideal for the reaction of primary amines with succinimidyl esters. To test the efficiency of the reaction and to judge the reactivity profile of the succinimidyl esters, three different ratios of dye to protein were used (2 (**33**), 6 (**34**), and 25 equiv. (**35**)). The reaction time in each of these three experiments was 15 min (as recommended by Molecular Probes Inc.) The 25 equivalent reaction was also repeated with a reaction time of 1 hour (**36**) to ascertain if reaction time was a factor in product formation.

For each experiment the 5-SFX dye (**31**) was dissolved in DMF and the appropriate volume added to the CV-N buffer solution and left at room temperature for the required time. After each reaction was complete the sample was chromatographed on a G25 size exclusion column (140 x 10 mm) eluting with phosphate buffered saline (PBS). On the column the sample separated into two coloured bands, the first eluting being light orange (~ 5 mL) and the second bright yellow. This was noted for all four reactions with the light orange band becoming more intense as the equivalents of dye and the reaction time were increased respectively. This was a positive sign that the reaction had taken place as the first band eluted would represent the fluorescent protein derivatives followed by the smaller unreacted and/or hydrolysed dyes. Each sample was submitted for LCMS analysis. The results are shown in **Table 4.1**.

Table 4.1 CV-N (**15**) reactions with 5-SFX (**31**) at pH 8.2

	Number of dye molecules attached (%)				
Conditions	**0**	**1**	**2**	**3**	**4**
5-SFX, 2eq, 15min (**33**)	42	48	10	0	0
5-SFX, 6eq, 15min (**34**)	7	46	37	10	0
5-SFX, 25eq, 15min (**35**)	0	15	47	30	8
5-SFX, 25eq, 60min (**36**)	0	0	9	47	44

The percentages expressed in **Table 4.1** are derived from the abundance of the relevant m/z ions in the mass spectra. This table is also displayed graphically below (**Figure 4.1**).

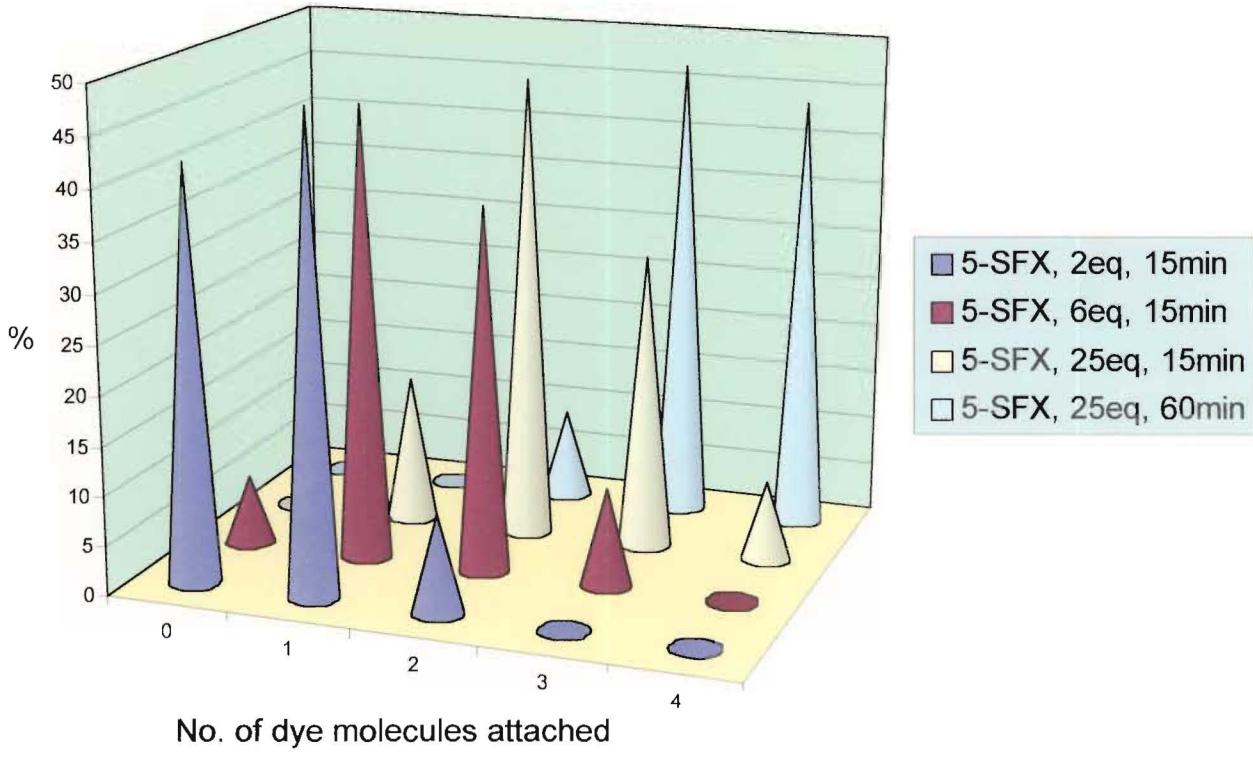


Figure 4.1 CV-N reactions with 5-SFX (**31**) at pH 8.2

The results show a clear trend with respect to the reactivity of the succinimidyl ester of 5-SFX. As the number of equivalents was increased from 2 to 6 to 25, the number of dye molecules attached increased significantly. The final data series (light blue) representing the extended reaction time of 1 hour also increased the average number of dye molecules with most CV-N molecules having 3 or 4 fluorescent tags. Interestingly, there was no sign of any conjugates beyond the attachment of four dye molecules. This could mean that there are one or two amines present in CV-N (**15**) that are not as accessible in terms of reactivity. CV-N (**15**) has 5 amines and an *N*-terminal, potentially allowing attachment of six dye molecules. **Figure 4.2** shows in three dimensions the positions of the lysine residues and the *N*-terminal derived from the published solution structure of CV-N.[85] As can be noted,

the side chains of the lysine residues seem to be free of any obstruction (not buried internally), although the visualisation does not highlight the rest of the amino acids' side chains (for clarity). However, even taking this factor into account, all amines appear to be pointed outwards, towards the surrounding solution and are therefore presumably available for reaction. Another, more appealing hypothesis is that once a dye molecule is attached to a lysine side chain, it may, due to steric interference, prevent a further succinimidyl group from reaching a neighbouring lysine amine. The *N*-terminal and adjacent lysine side chain amines are a mere 5.1 Å apart while the two lysine side chains to their left in **Figure 4.2** are relatively close at 13.7 Å. It will ultimately take an enzyme digest analysis to determine where exactly the dye molecules are attached and even then, it may still prove difficult to distinguish a clear relative reactivity of the amines.

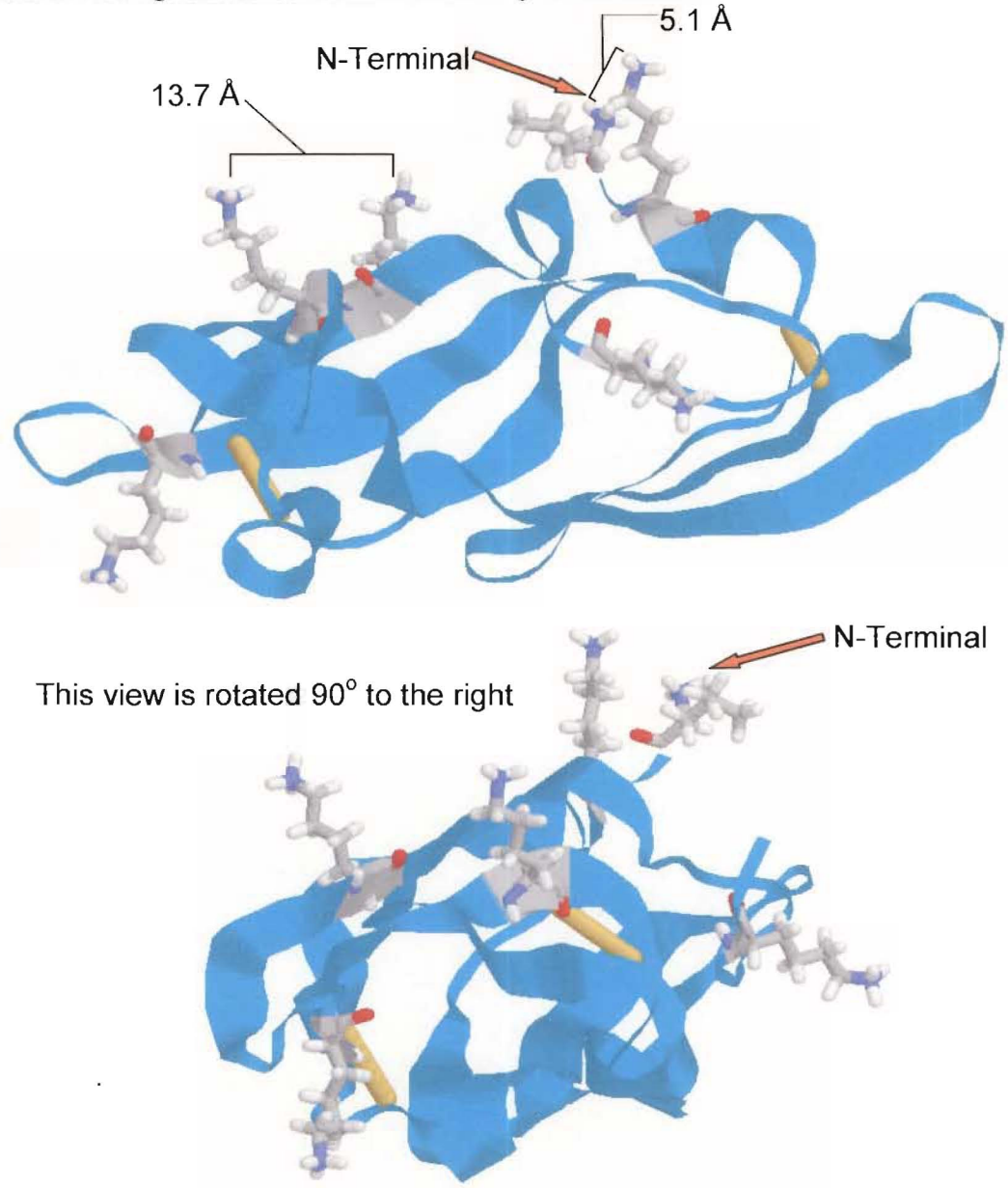


Figure 4.2 CV-N with lysine amines and *N*-terminal highlighted

In summary, this first series of reactions indicated that CV-N (**15**) was capable of reacting with succinimidyl esters in a concentration and time dependent manner with up to four dye molecules able to be attached.

4.2.2.2 Reactions at pH 7.2

The optimum pH for the reaction of succinimidyl esters was suggested as ~pH 8.0 to 9.5, however, literature[149] had suggested the possibility of *N*-terminus amine selectivity if the pH of reaction was lowered to around 7. This is due to the differential pKa of an *N*-terminus (~7) amine *versus* a lysine side chain amine (~10.5). To determine if this may be the case, and to also test the reactivity of the succinimidyl esters at near neutral pH, a series of reactions was also carried out at pH 7.2.

The reaction conditions were exactly the same as those in section **4.2.2.1** with 2 (**37**), 6 (**38**), and 25 equiv. (**39**) of dye reacted for 15 min and 25 equiv. reacted for 60 min (**40**). The only difference was the change of buffer, from $NaHCO_3$ to PBS. After the reactions were completed, the samples were again chromatographed on a small G25 column, but this time it was much more difficult to determine coloured bands representing the protein-dye conjugates. As elution times were consistent for all reactions of this type, fractions were still collected at the appropriate times for separation of 'protein' material and 'small molecule' material even if they could not be determined visually. The results are shown in **Table 4.2** and **Figure 4.3** below and indicate a significant difference from the series of reactions at pH 8.2. Only 3% of the CV-N (**15**) was converted to the single dye form at 2 equiv. with still only 21% even at 25 equiv. The final reaction with 25 equiv. over a 1 hour reaction period showed 47% of single dye and 8% of two dye conjugates. A comparison between **Figure 4.1** and **Figure 4.3** shows this decrease in reactivity from pH 8.2 to 7.2.

Table 4.2 CV-N (**15**) reactions with 5-SFX (**31**) at pH 7.2

	Number of dye molecules attached (%)				
Conditions	**0**	**1**	**2**	**3**	**4**
5-SFX, 2eq, 15min (**37**)	97	3	0	0	0
5-SFX, 6eq, 15min (**38**)	89	11	0	0	0
5-SFX, 25eq, 15min (**39**)	79	21	0	0	0
5-SFX, 25eq, 60min (**40**)	45	47	8	0	0

It was difficult to tell from this series of reactions whether the data point towards an attachment of dye at primarily a single position (the *N*-terminal), or whether the preponderance of only a single attachment meant the reaction had slowed due to the non-optimal pH. One of the reasons for this series of experiments was to determine if *N*-terminal selectivity did occur in order to justify the original strategy of attaching the terminal tripeptide with toxin attached to the CV-N(-3N) as discussed in section **2.2**.

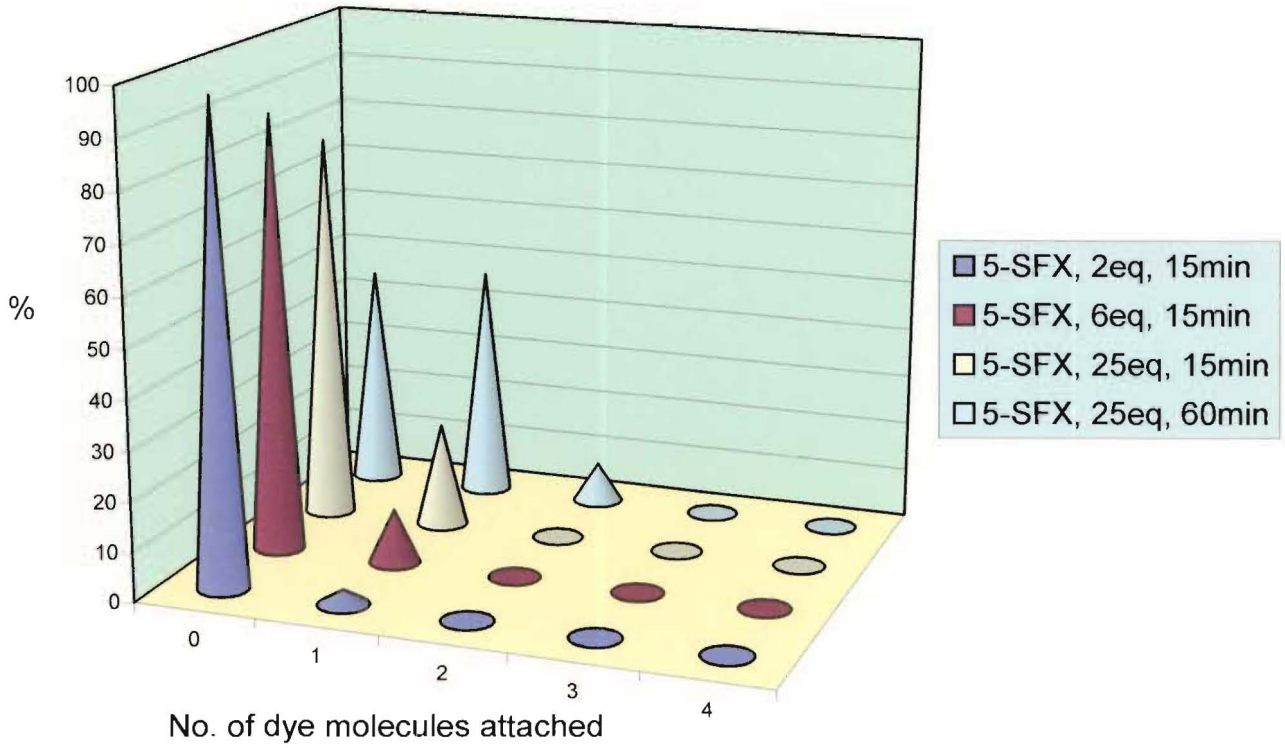


Figure 4.3 CV-N reactions with 5-SFX (**31**) at pH 7.2

These results, and those to come with the BODIPY dye, were not as conclusive as would be needed to validate that particular strategy which is one of the reasons it was discontinued. The lack of general reaction for these four experiments demonstrated how crucial the pH was in the reactivity of lysine primary amines towards succinimidyl esters.

4.2.3 BODIPY FLX-SE Reactions with CV-N

4.2.3.1 Reactions at pH 8.2

In order to contrast the reactivity profile of the succinimidyl esters present in the 5-SFX (**31**) compound, a separate dye, BODIPY FLX-SE (**32**), was also reacted in the same way. The experimental protocols were identical to those explained in section **4.2.2.1** with only the identity of the dye changing. After reaction, the colours of the eluted bands from the G25 column were orange for the protein-dye conjugates and a light green colour for unreacted dye. For the two reactions with 25 equiv. of the BODIPY dye (**32**) for a 15 (**41**) and 60 min (**42**) reaction time, reaction appeared to take place with coloured bands eluting as usual. However, mass spectral analysis by LCMS failed to show any sign of CV-N-dye conjugates or indeed, native CV-N (**15**). The coloured band eluting at the correct time for these reactions suggested that the conjugation had taken place with CV-N, so it was therefore puzzling to be unable to see the results by LCMS. A sample of these two fractions was further purified by loading onto a small C4 cartridge, washing with water to remove salts and buffers and then eluting with MeOH. These were then dried under a stream of N_2 and resubmitted for LCMS analysis. Again, no sign of CV-N (**15**) or a conjugated derivative was observed. A possible explanation is that these reactions sufficiently 'loaded' the CV-N with multiple dye attachments and this prevented protonation and ionisation in the mass spectrometer, although this did not take place for the 5-SFX reactions with high loading. As the remaining reactions worked well it was not, however, considered to be a major obstacle. The results for the first two reactions with BODIPY using 2 equiv.(**43**) and 6 equiv. (**44**) are shown in **Table 4.3** and **Figure 4.4**.

Table 4.3 CV-N (**15**) reactions with BODIPY FLX-SE (**32**) at pH 8.2

	Number of dye molecules attached (%)				
Conditions	**0**	**1**	**2**	**3**	**4**
BODIPY FLX, 2eq, 15min (**43**)	75	25	0	0	0
BODIPY FLX, 6eq, 15min (**44**)	40	60	0	0	0

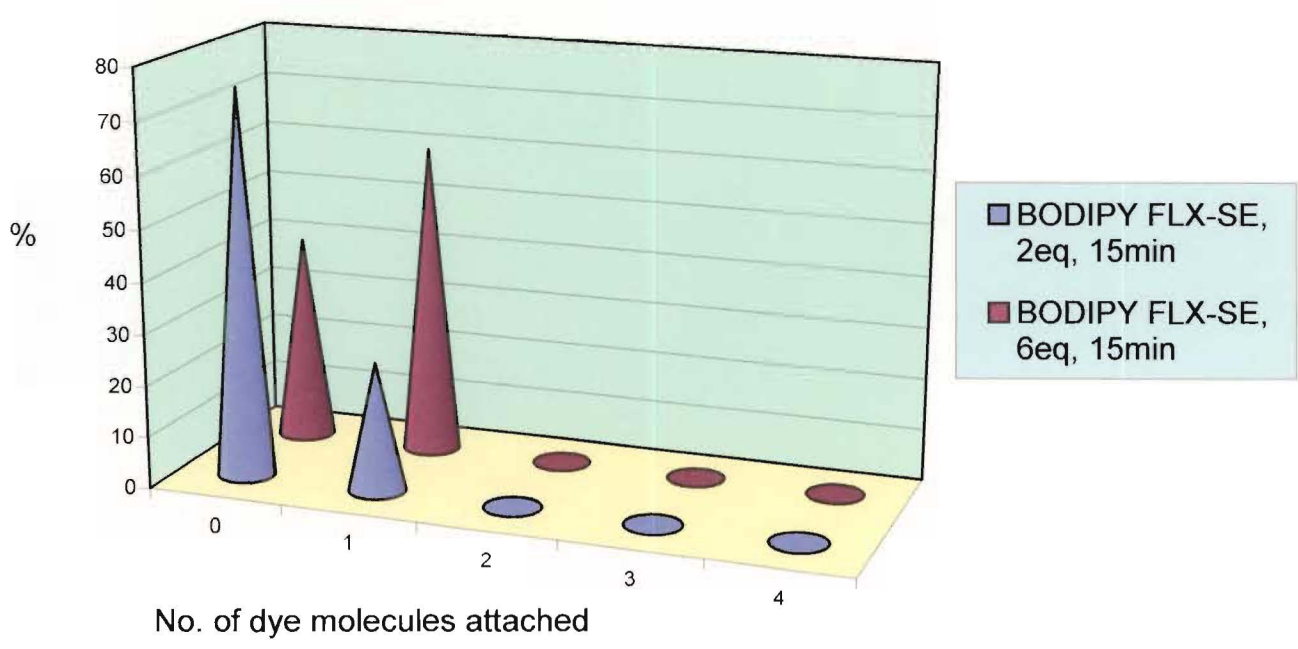


Figure 4.4 CV-N reactions with BODIPY FLX-SE (**32**) at pH 8.2

It is difficult to precisely define a trend with only two representative points but it was clear that as the number of equivalents increased from 2 to 6, the percentage of conjugates with one dye molecule attached increased. Compared to the results from the 5-SFX dye (**31**) (**Table 4.1**) the BODIPY dye (**32**) does not seem to be as reactive at this pH.

4.2.3.2 Reactions at pH 7.2

This series followed the protocol for the above reactions except at the lower pH of 7.2. There were no problems with the 25 equiv. reactions this time. The results are shown in **Table 4.4** and in **Figure 4.5**. Both the 2 (**45**) and 6 equiv. (**46**) experiments showed no modification to CV-N (**15**) whatsoever with moderate rates of attachment for the dual 25 equiv. (**47**, **48**) reactions. A comparison of the two dyes shows, that like the reactions at pH 8.2, the BODIPY dye (**32**) appears to react more slowly than the 5-SFX (**31**). The results are again inconclusive as to *N*-terminal selectivity of the succinimidyl ester at the lower pH.

Table 4.4 CV-N (**15**) reactions with BODIPY FLX-SE (**32**) at pH 7.2

	Number of dye molecules attached (%)				
Conditions	**0**	**1**	**2**	**3**	**4**
BODIPY FLX, 2eq, 15min (**45**)	100	0	0	0	0
BODIPY FLX, 6eq, 15min (**46**)	100	0	0	0	0
BODIPY FLX, 25eq, 15min (**47**)	82	17	0	0	0
BODIPY FLX, 25eq, 60min (**48**)	63	37	0	0	0

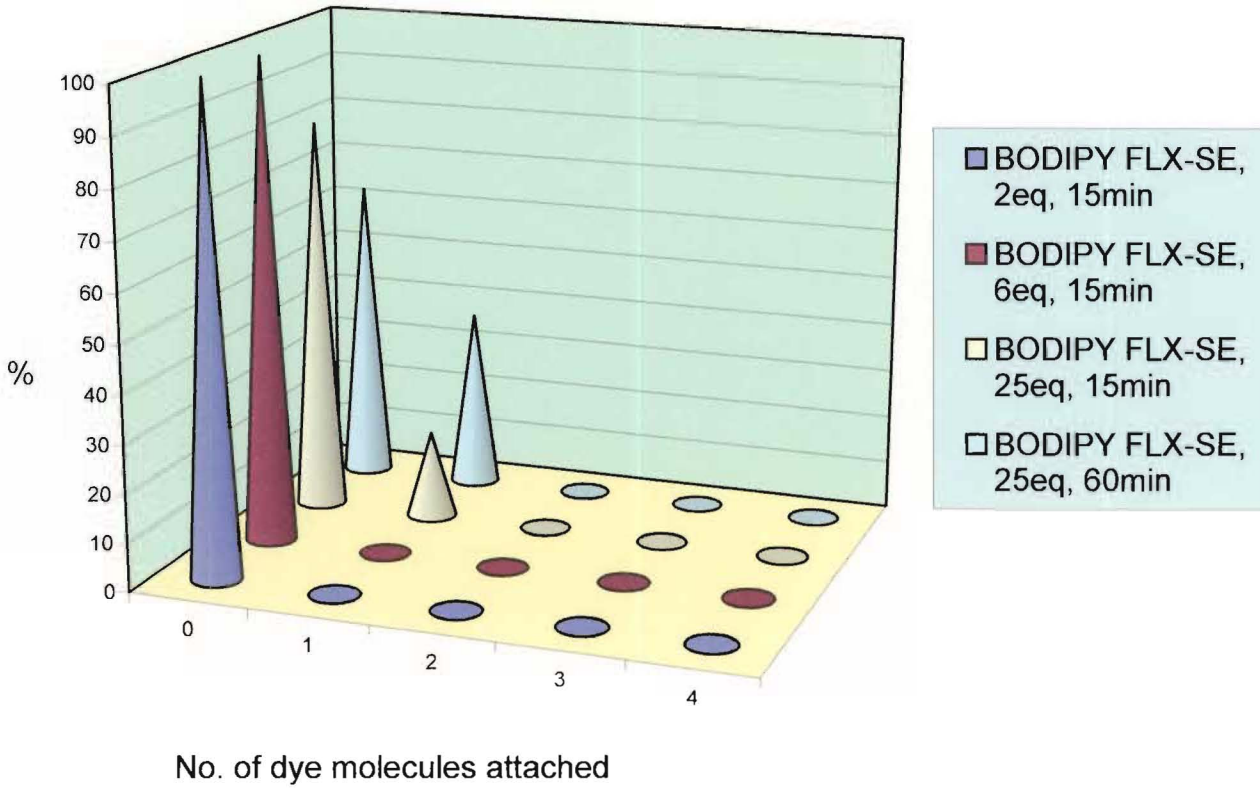


Figure 4.5 CV-N reactions with BODIPY FLX-SE (**32**)at pH 7.2

4.2.4 Conclusions Reached from the Dye Reactions

The first premise from this series of reactions was to confirm that the succinimidyl esters do in fact react with CV-N (**15**) forming stable amide bonds through the lysine side chain and/or *N*-terminal amines. These reactions clearly showed extensive reaction taking place in a concentration and time-dependent manner. The 5-SFX dye (**31**) appeared more reactive than its BODIPY counterpart (**32**) and both reacted more slowly at pH 7.2 than at 8.2. Whether the lower rate of reaction at the lower pH was due to *N*-terminal selectivity or merely a consequence of the non-optimum pH was difficult to tell, and may be a combination of both factors. Further analysis of these samples to determine their binding affinities to gp120 and anti-HIV properties was also undertaken (section **6.3**). In short, the results from the testing of these conjugates using the ELISA and XTT assays showed that the dye molecules did not interfere with binding of CV-N to gp120. Therefore, the strategy to attach conjugates to CV-N (**15**) in this way could be continued (as follows).

4.3 Biolinker-Coumarin Constructs

While the reactions involving the fluorescent dye conjugates in section **4.2** showed that succinimidyl esters would react with amine groups in CV-N (**15**), there remained the question of whether or not the rest of the synthetic strategy outlined in section **2.4** and **Figure 2.3** would work. There needed to be a systematic approach to joining the succinimidyl ester, biolinker, and toxin together in a way to maximise the yields of reaction while preserving the bioactivity of the toxin. A fluorescent coumarin derivative was chosen as a substitute for the toxin as it was commercially available, had a primary amine, and had the advantage of fluorescence for ease of purification and identification in further testing.

4.3.1 Peptide Biolinker

4.3.1.1 Peptide Bond Formation and Solid Phase Peptide Synthesis

Formation of a peptide bond requires that the carboxyl group of an amino acid be activated (X) while the amino group (Y) and non-participating carboxyl of the other amino acid (Y') be protected[153] as shown in **Figure 4.6**. It is also necessary for some side chains of particular amino acids (e.g. lysine) to be protected so that they do not interfere with the coupling process. During this coupling the chiral integrity of the amino acids must also be preserved. In an ideal situation the protecting groups must be orthogonal to one another which means that they must be easily removable without simultaneous loss of the others, and *vice versa*. There is a wide selection of protecting groups available from which a

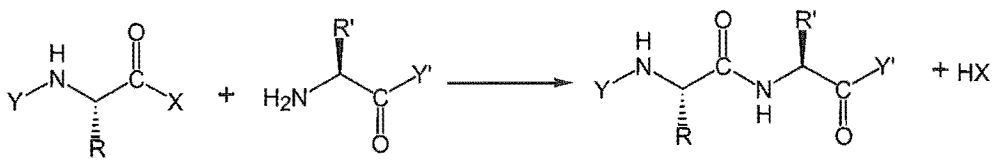


Figure 4.6 Model peptide coupling

strategy needs to be formulated to accomplish the specific peptide goal.

The most efficient method for synthesising peptides is by using 'Solid Phase Peptide Synthesis' (SPPS) techniques. The synthesis is carried out from the *C* to *N*-terminal of the peptide. This involves sequential addition of α-amino and side chain protected amino acid residues to an insoluble polymeric support (resin bead). After removal of the *N*-α protecting group, the next protected amino acid is joined to the peptide by a coupling reagent. This process is repeated until the desired peptide length is reached, at which point the product is cleaved from the resin. The two most common protecting groups used for SPPS are the acid labile Boc group and the base labile Fmoc group. Of the two, Fmoc was chosen as more suitable due to the hazardous use of HF for the final cleavage of the Boc strategy and also concerns about the alteration of sensitive peptide bonds with the repetitive acid treatment used for Boc deprotection. Fmoc SPPS[154] utilises piperidine (mild base) to remove the protecting group and TFA to cleave the peptide from the resin.

4.3.1.2 Synthesis of Fmoc-Gly-Phe-Leu-Gly-OH

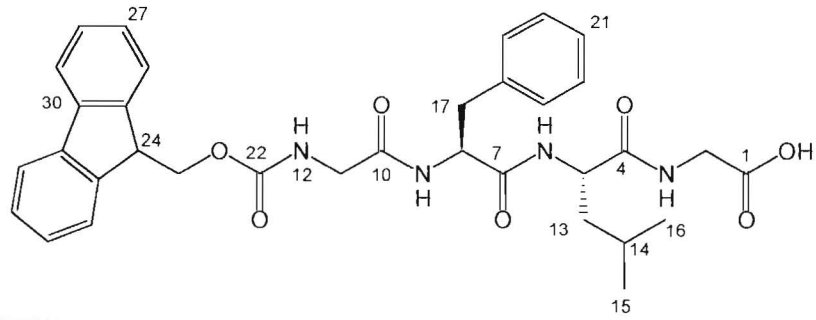
The aim of this experiment was to produce the tetrapeptide Fmoc-Gly-Phe-Leu-Gly-OH for use with the various synthetic strategies utilising this biolinker. The glassware apparatus used for SPPS allowed the resin to be bubbled in solution with N_2 over a frit and subsequently switched to vacuum to allow drainage and rinsing at the various stages of the synthesis. Fmoc-Gly-Wang resin (3 g, 0.897 meq/g) was bubbled in DMF for 30 min to pre-swell the beads allowing greater access for reagents in solution. The DMF was drained and a 20% piperidine/DMF solution (30 mL) added, with bubbling, to deprotect the glycine residue. After 10 min the reaction mixture was drained and washed with DMF and then IPA. At this stage a couple of resin beads were removed from the apparatus, washed with EtOH and submitted for Kaiser test analysis (discussed in section **8.1**) to determine the presence of free amine. A positive test was obtained (blue resin beads) and the synthesis was allowed to continue. Once washed thoroughly, a DMF solution containing Fmoc-Leu (2 equiv.), HBTU (coupling reagent, 2 equiv.), HOBt (to prevent racemisation, 2 equiv.),

DIPEA (base, 4 equiv.) was added and bubbled for 60 min. After rinsing with DMF and again performing a Kaiser test (this time negative) the resin was ready for the procedure to begin again from the deprotection of the Fmoc group to the coupling of the remaining two amino acids (Fmoc-Phe and then Fmoc-Gly). After the final amino acid was coupled to the peptide the Fmoc-protected biolinker was left drying overnight in the presence of P_2O_5 under vacuum. The peptide was now ready for cleavage from the resin with 95% TFA, 2.5% H_2O, and 2.5% triethylsilane (TES, scavenger). After 20 min of bubbling, the solution was drained, rinsed with fresh TFA and collected into a flask. This solution was then concentrated on a rotary evaporator. Water was added to precipitate the peptide out of the acidic solution. The solution was next evaporated to remove more TFA and this process repeated with fresh additions of water each time to dilute out the acid solution. Water was added one last time and the resulting precipitate filtered on a sintered glass funnel. The Fmoc-Gly-Phe-Leu-Gly-OH (**49**) solid was then dried and weighed. (1.17 g, 71%). Small amounts of the product were detected in the aqueous wash by FABMS, but the purity was very low and made recovery unworthwhile.

4.3.1.3 Characterisation of Fmoc-Gly-Phe-Leu-Gly-OH

The ^1H and ^{13}C NMR spectra of **49** were readily assignable with reference to COSY, HSQC and HMBC 2D NMR experiments. The chemical shifts for the ^1H and ^{13}C NMR spectra are shown in **Table 4.5** with the structure and major HMBC correlations shown in **Figure 4.7**. The unique binding of amino acids in peptides allows 2D NMR techniques, particularly HMBC, to bridge across the amide bond of the peptide allowing a complete unambiguous assignment of structure.

A good starting point for the structural elucidation was to use the readily assigned methyls of C15 and C16 at 0.79 ppm. HSQC correlations confirmed the respective ^{13}C resonances, while HMBC and COSY correlations from the methyl protons allowed assignment of C14 and C13. A COSY correlation from the H13 resonances allowed assignment of the leucine α-proton, H5. From H5, a COSY correlation to the amide NH6 gave full assignment of all



Fmoc-Gly-Phe-Leu-Gly-OH (**49**)

Table 4.5 ^{1}H and ^{13}C NMR data for Fmoc-Gly-Phe-Leu-Gly-OH (**49**)

	^{1}H δ (ppm)[a]	^{13}C δ (ppm)[b]		^{1}H δ (ppm)[a]	^{13}C δ (ppm)[b]
1	-	171.2	**16**	0.79 (d, *J*=5.9 Hz)	22.6 or 21.0
2	3.79 (m)	40.7	**17**	2.97 (m)	37.2
3NH	7.14[c]	-	**18**	-	135.8
4	-	172.6	**19**	7.09 (m)	128.9
5	4.33 (m)	51.4	**20**	7.12 (m)	128.3
6NH	7.64 (m)	-	**21**	7.09 (m)	126.7
7	-	171.2	**22**	-	157.1
8	4.53 (m)	54.3	**23**	4.25 (m)	67.0
9NH	7.4 (d, *J*=7.3 Hz)	-	**24**	4.12[c]	46.7
10	-	170.2	**25**	-	143.4
11	3.68 (m)	43.9	**26**	7.52 (d *J*=6.8 Hz)	124.7
12NH	6.61 (m)	-	**27**	7.23 (t, *J*=6.8 and 7.5 Hz)	126.8
13a	1.44 or 1.53 (m)	39.9	**28**	7.32 (t, *J*=7.5 and 7.5 Hz)	127.5
13b	1.44 or 1.53 (m)	39.9	**29**	7.69 (d *J*=7.5 Hz)	119.7
14	1.44 (m)	24.2	**30**	-	141.0
15	0.79 (d, *J*=5.9 Hz)	22.6 or 21.0			

[a] Data recorded at 23°C in $CDCl_3$ at 300 MHz with chemical shifts in ppm and referenced to $CHCl_3$, δ_H 7.26

[b] Data recorded at 23°C in CDCl3 at 75 MHz with chemical shifts in ppm and referenced to CDCl3, δ_C 49.3

[c] peak obscured therefore no multiplicity available

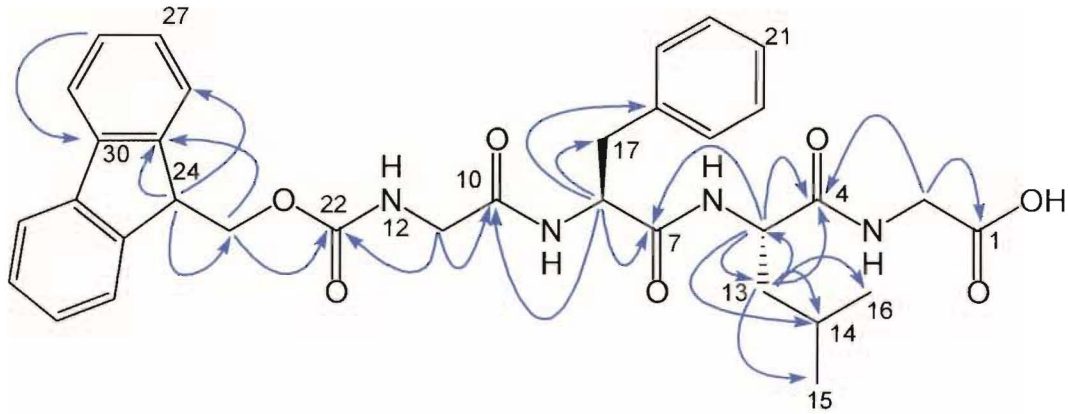


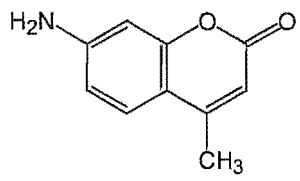
Figure 4.7 Important HMBC correlations for Fmoc-Gly-Phe-Leu-Gly-OH (**49**)

of the protons of the leucine residue. H5 also showed two HMBC correlations to carbons in the carbonyl region of the ^{13}C spectra but only one of these two was seen as an HMBC correlation from H13. This therefore could be assigned as C4 as only the leucine carbonyl was able to show HMBC correlations from the H5 and H13 protons. Another HMBC correlation to this leucine carbonyl (C4) allowed assignment of the adjacent glycine α-protons H2 and the amide 3NH. All HSQC correlations in the spectra were present therefore allowing immediate correlation of ^{1}H chemical shifts to their ^{13}C counterparts. The second HMBC correlation from H5 to a carbonyl resonance allowed assignment of C7 although this turned out to have the same ^{13}C chemical shift as C1. HMBC correlations to C7 from H17 and H8 respectively gave their chemical shifts, confirmed by COSY correlations from H17 to H8 and from H8 to NH9. Access into the aromatic ring of phenylalanine was gained through HMBC correlations from H8 and H17. The chemical shifts of H28, H29 and H30 were overlapped in the ^{1}H spectrum but HMBC analysis elucidated their distinctive ^{13}C resonances. The carbonyl, C10, was found from an HMBC correlation from H8 which gave access to the glycine adjacent to the Fmoc group. Another HMBC to this carbonyl identified the α-protons H11 with a COSY from H11 giving NH12. Linkage of the Fmoc group could be confirmed from two HMBC correlations to the carbonyl C22 from both H11 and H23. From H23 and C23 it was possible to follow the HMBC correlations around the remainder of the Fmoc group with HSQC correlations elucidating the respective proton resonances.

The molecular formula for **49**, $C_{34}H_{38}O_7N_4$ was obtained by HRFABMS as the sodiated adduct of mass 637.27 Da (2.93 ppm). An analytical isocratic HPLC injection (C18, 80% MeOH/H_2O (0.05% TFA)) yielded a single peak only, at 7.6 min retention time with a characteristic UV chromophore representing primarily the Fmoc group of the tetrapeptide.

4.3.2 Fmoc-Gly-Phe-Leu-Gly-Coumarin

The structure of 7-amino-4-methylcoumarin (**50**) is shown below. The direct attachment of the amine to an aromatic ring reduces the nucleophilicity of this amine and led to initial problems in coupling the coumarin moiety to the *C*-terminal of **49**. The delocalising effect of the aromatic ring on the nucleophilic potential of the coumarin amine meant it could not be coupled *via* a succinimidyl ester as planned for the halichondrin derivative. However, a search of the literature revealed a method for coupling the coumarin amine to the carboxyl of the tetrapeptide using $POCl_3$ in pyridine.[155] These conditions were obviously too harsh for the toxin to survive but this first reaction would be accomplished *via* a succinimidyl ester when incorporating a toxin molecule.



7-amino-4-methylcoumarin (**50**)

4.3.2.1 Synthesis of Fmoc-Gly-Phe-Leu-Gly-Coumarin

To a solution of **49** in freshly distilled pyridine was added **50** (1.1 equiv.) along with $POCl_3$ (1.1 equiv.) (**Figure 4.8**). After stirring for 1 hour the solution was dried under vacuum. HPLC analysis on an analytical C18 column (80% MeOH/H_2O (0.05% TFA)) allowed identification of the product from the change in both retention time and UV chromophore from the starting material. This elution profile was then transferred to a preparative C18 column and the resulting peaks collected with identification of the product at a retention time of 27 min. The Fmoc-Gly-Phe-Leu-Gly-coumarin (**51**) product was obtained in a yield of 67%.

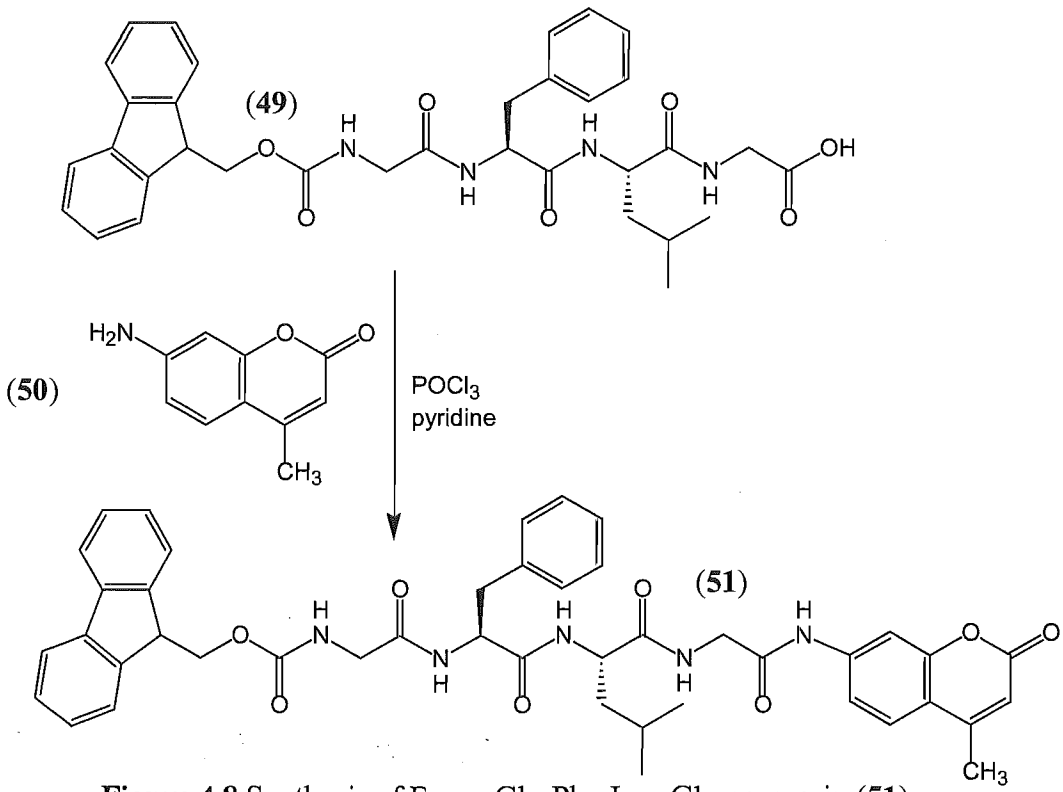
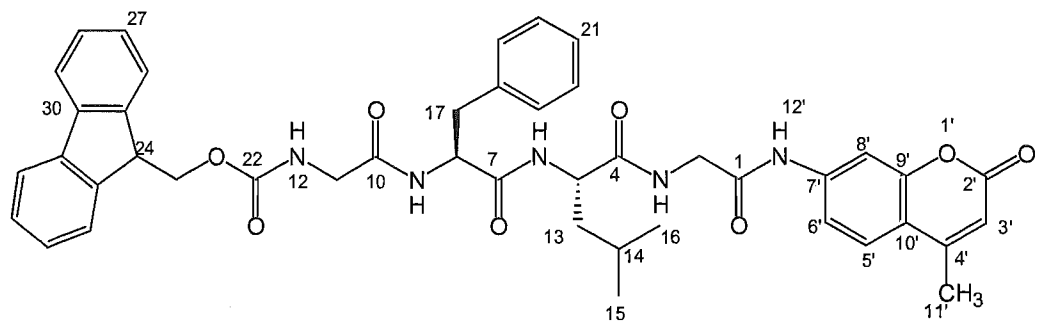


Figure 4.8 Synthesis of Fmoc-Gly-Phe-Leu-Gly-coumarin (**51**)

4.3.2.2 Characterisation of Fmoc-Gly-Phe-Leu-Gly-Coumarin

Compound **51** would not dissolve in $CDCl_3$ for NMR analysis. CD_3OD was therefore used, but resulted in a loss of the amide signals after a short period of time. The change of solvent also made the direct comparison of the NMR spectra of the tetrapeptide starting material (**49**) and **51** more difficult. However, the full structural elucidation was completed through use of COSY, HSQC and HMBC experiments as for the starting material (**Table 4.6**). The distinctive chemical shifts of the H3' and H11' were significant in identifying the other components of the coumarin section of the molecule. HMBC correlations were again significant in assigning the ^{13}C NMR spectrum. High resolution ESIMS confirmed the mass of 772.3341 Da (MH^+, 0.7 ppm) for $C_{44}H_{46}N_5O_8$.



Fmoc-Gly-Phe-Leu-Gly-coumarin (**51**)

Table 4.6 ^{1}H and ^{13}C NMR data for Fmoc-Gly-Phe-Leu-Gly-coumarin (**51**)

	^{1}H δ (ppm)[a]	^{13}C δ (ppm)[b]		^{1}H δ (ppm)[a]	^{13}C δ (ppm)[b]
1	-	169.4	**21**	7.1 (m)	127.7
2a	3.85 (d, *J*=17.1 Hz)	44.0	**22**	-	158.4
2b	4.06 (d, *J*=17.1 Hz)		**23**	4.25 (m)	68.1
3NH	[c]	-	**24**	4.14 (m)	47.6
4	-	174.3	**25**	-	144.3
5	4.26 (m)	53.8	**26**	7.51 (m)	125.9
6NH	[c]	-	**27**	7.22 (m)	128.1
7	-	173.3	**28**	7.34 (m)	128.5
8	4.57 (t, *J*=5.8 Hz)	55.5	**29**	7.70 (m)	120.6
9NH	[c]	-	**30**	-	141.9
10	-	172.0	**1'**	-	-
11	3.76 (bs)	45.0	**2'**	-	163
12NH	[c]	-	**3'**	6.15 (s)	113.4
13a	1.51 or 1.61 (m)	40.2	**4'**	-	154.4
13b	1.51 or 1.61 (m)	40.2	**5'**	7.63 (m)	127.6
14	1.51 (m)	25.3	**6'**	7.65 (m)	125.7
15	0.85 or 0.87 (d, *J*=5.9 Hz)	21.9 or 23.4	**7'**	-	144.4
16	0.85 or 0.87 (d, *J*=5.9 Hz)	21.9 or 23.4	**8'**	7.66 (s)	108.0
17a	3.01 (dd, *J*=7.1 and 14.0 Hz)	37.7	**9'**	-	154.5
17b	3.12 (dd, *J*=6.1 and 14.0 Hz)		**10'**	-	116.8
18	-	136.6	**11'**	2.37 (s)	19.0
19	7.09 (m)	129.8	**12'NH**	[c]	-
20	7.12 (m)	129.2			

[a] Data recorded at 23°C in CD_3OD at 300 MHz with chemical shifts in ppm and referenced to CD_2HOD, δ_H 3.31

[b] Data recorded at 23°C in CD_3OD at 75 MHz with chemical shifts in ppm and referenced to CD_3OD, δ_C 49.3

[c] The amides were only partially observed in the ^{1}H spectrum but had exchanged before characterisation from other experiments could take place

4.3.2.3 Fmoc Removal to form H_2N-Gly-Phe-Leu-Gly-Coumarin

The suggested solvent for removal of the Fmoc group is DMF which is much less volatile and hence, much more difficult to remove than CH_3CN. Mild conditions for solvent removal would be needed when the toxin was present. Therefore, this reaction was attempted in CH_3CN (**Figure 4.9**) and was, therefore a test of whether the Fmoc group could be removed under these conditions. The Fmoc-Gly-Phe-Leu-Gly-coumarin (**51**) was dissolved in 20% piperidine/CH_3CN and left for 30 min. The solvent was removed and the

sample dried under vacuum. The solid was dissolved in DCM and chromatographed on a small silica column. Three fractions were collected, eluting DCM, EtOAc and finally MeOH. The Fmoc-adduct (**52**), eluted in the EtOAc fraction while the product (**53**) was not eluted until the MeOH fraction.

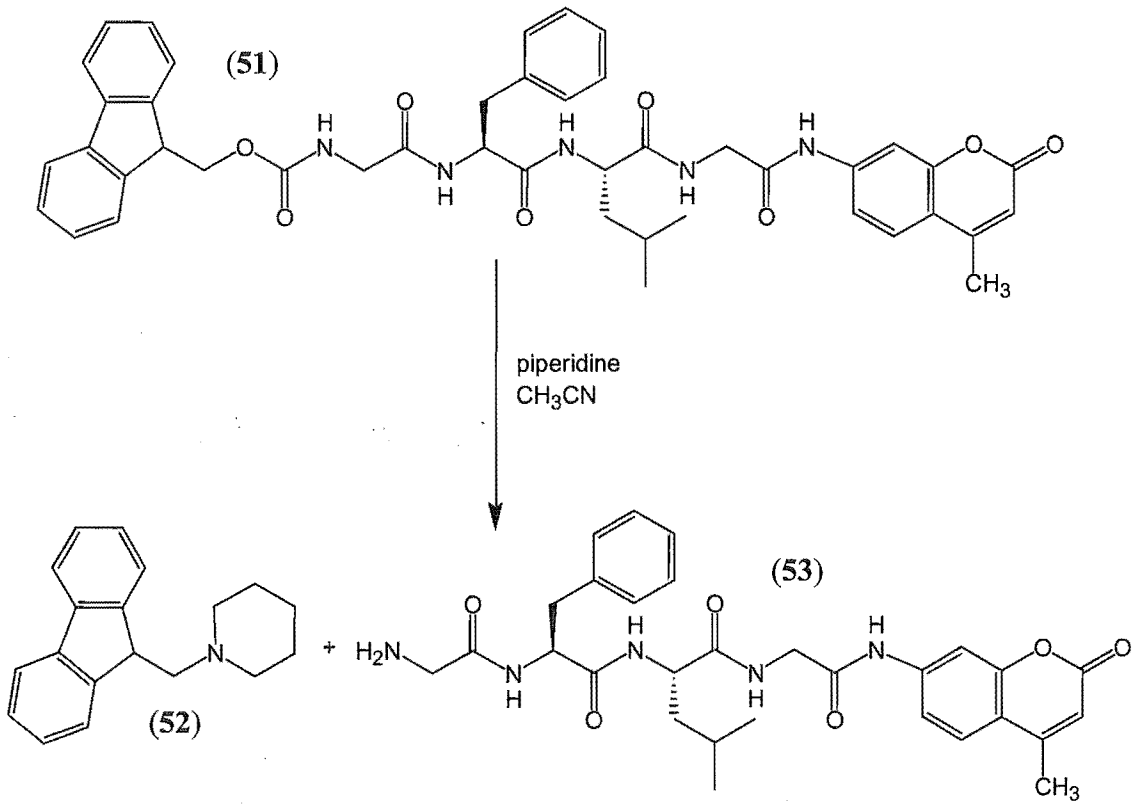


Figure 4.9 Removal of Fmoc group from Fmoc-Gly-Phe-Leu-Gly-coumarin (**51**)

Characterisation of H_2N-Gly-Phe-Leu-Gly-coumarin (**53**) was by direct comparison of the ^1H NMR spectrum with **51** and by ESIMS analysis. The removal of the Fmoc group made the characterisation much easier as the many aromatic resonances of Fmoc were removed from the spectrum. There was also a significant downfield shift of the *N*-terminal glycine α-proton resonances. This was expected from conversion of the adjacent amide to an amine.

4.3.3 Succinimide Activation of H_2N-Gly-Phe-Leu-Gly-Coumarin

The initial strategy at this stage was to activate the free terminal amine with a succinimidyl ester in order to react it with CV-N (**15**). The reagent N,N'-disuccinimidyl carbonate (DSC, **54**)[143] is capable of converting amines to succinimidyl esters, as shown in **Figure 4.10**, forming a carbamate or urethane bond. Once the carbamate ester reacts with CV-N (**15**) a stable urea bond would be formed.

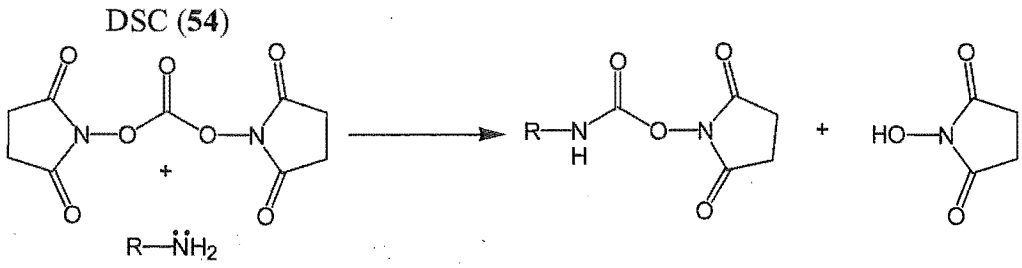


Figure 4.10 Activation of amines by DSC (**54**)

The reaction between DSC (**54**) and H_2N-Gly-Phe-Leu-Gly-coumarin (**53**) was attempted many times with varying solvents and conditions. Although the correct product was seen by ^{1}H NMR spectroscopy and ESIMS analysis of the crude samples, purification predominantly led to a compound with a mass 26 Da higher than that for the amine, but 115 Da less than that required for the succinimidyl ester. The presence of the succinimidyl ester in the crude sample was relatively easy to identify in the ^{1}H NMR spectra (singlet of integral 4, at ~2.8 ppm). However, the intensity of the corresponding resonance decreased no matter what method was attempted for purification (e.g. column chromatography or HPLC). The other striking features of the resulting ^{1}H NMR spectrum after 'purification' when compared to the starting material amine (**53**) was the downfield shift of the α-proton of phenylalanine by 0.50 ppm and the β-protons by 0.16 ppm respectively. This data, in combination with a shift in the *N*-terminal glycine α-protons of 0.22 ppm provided evidence for a dramatic change in the molecule. Possible adducts of the succinimidyl esters were investigated that matched both the NMR data and the ESIMS results. One proposal examined was that the phenylalanine amide had attacked the activated succinimidyl ester and produced a five membered imidazolidine ring as in **Figure 4.11**.

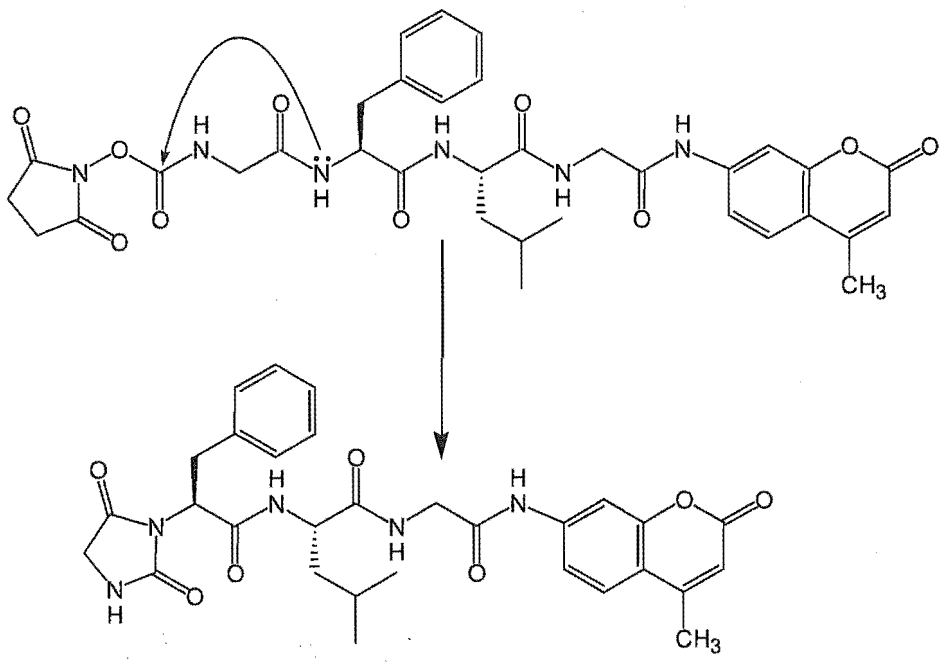


Figure 4.11 Postulated reaction of succinimidyl ester of biolinker-coumarin conjugate

This would explain both the increase in mass from the parent amine by 26 Da and also the downfield ^{1}H NMR shift in the phenylalanine and glycine protons. An HMBC was attempted on the material but the sample was not completely pure and there was not enough material to resolve the signals representing the α-proton of phenylalanine correlating to three distinct carbonyls or from the *N*-terminal glycine α-proton to two distinct carbonyls. This would have gone a long way to proving the postulated adduct was indeed present.

4.3.3.1 Chain Extension Reaction with Succinic Anhydride

To circumvent this problem, **53** was reacted with succinic anhydride (1.2 equiv.) and DMAP (1 equiv.) in pyridine to extend the chain and convert the amine to a carboxyl in the hope that this would prevent the ring forming reaction (**Figure 4.12**). The succinic acid adduct was duly purified (HPLC, C18, 40% CH_3CN/H_2O (0.05% TFA)) and again reacted in the same way as above, with DSC (**54**) and pyridine in CH_3CN. LCMS analysis using

the same elution conditions showed formation of the succinimidyl ester of the succinic-biolinker-coumarin. However, as the reaction progressed a new peak began to be visible after some time, representing the product shown in **Figure 4.12**. The succinimidyl ester did form, but as before, it subsequently reacted internally displacing the ester in favour of this time, a succinimide ring. This was further proof of the postulated 'ring closing' reaction which made this particular reaction pathway unfavourable. The results of these reactions were unacceptable in terms of transferring this strategy to incorporate the toxin molecule, therefore the strategy was modified (section **4.3.4**). Precedents for this type of intramolecular reaction were found in the literature particularly with the succinic spacer present, although the leaving group was usually a drug (not NHS as in this case).[156-158]

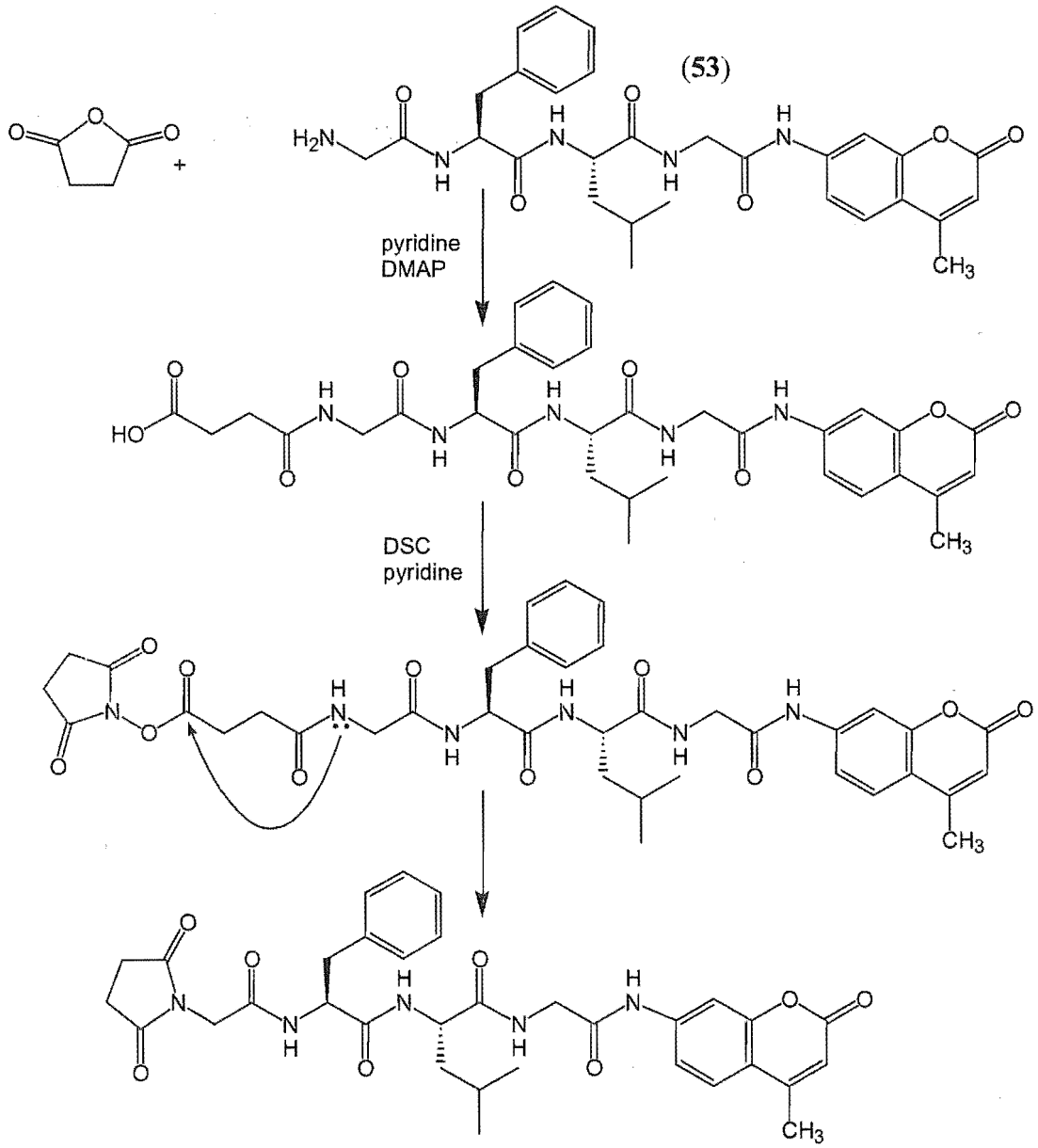


Figure 4.12 Succinic anhydride reaction and subsequent activation and 'ring closure'

4.3.4 Adipic Anhydride Reaction with H_2N-Gly-Phe-Leu-Gly-Coumarin

An analysis of the two previous reactions demonstrated that in order for the 'ring closing' reaction to occur the amide needed to be four atoms away from the activated carbonyl. Therefore, an obvious solution was to extend the chain by greater than four atoms. A chain length of 3 CH_2's would still afford the possibility of a six-membered ring, therefore adipic anhydride (**55**) was chosen to react with **53** as any ring closing reaction would have to form a less favoured 7-membered ring.

4.3.4.1 Adipic Anhydride Synthesis

Adipic acid (5 gm) and acetic anhydride (9.7 mL) were refluxed under argon for two hours at 160^0C.[159] The excess acetic anhydride and the acetic acid formed were distilled off at atmospheric pressure. The remaining liquid solidified on cooling and was distilled under reduced pressure on a Kugelrohr (Büchi GKR-50). The resulting liquid, which solidified at room temperature was analysed by ^{1}H NMR spectroscopy. The two peaks in the ^{1}H NMR spectrum matched closely to published data and an EIMS analysis confirmed the correct mass (128 Da).

4.3.4.2 Adipic Anhydride Reaction with H_2N-Gly-Phe-Leu-Gly-Coumarin

This approach required the terminal amine of the biolinker-coumarin to attack the reactive adipic anhydride (**55**) to form a pendant chain at the end of the biolinker (**Figure 4.13**). The H_2N-Gly-Phe-Leu-Gly-coumarin (**53**, 5.6 mg) was dissolved in CH_3CN and adipic anhydride (**55**, 1.1 equiv.) added and left for 24 hours. The reaction solvent was removed and the remaining solid dried under vacuum. ESIMS analysis of the crude residue showed the correct mass of 678 Da (MH^+). Preparative HPLC (C18, 50% CH_3CN/H_2O (0.05% TFA)) eluted a compound (**56**) at a retention time of 11.6 min with the correct UV chromophore (coumarin like). The ^{1}H NMR spectrum was readily assignable, with a

COSY analysis allowing assignment of the CH_2's of the adipic acid group. ESIMS analysis of the purified sample confirmed the molecular formula of the product.

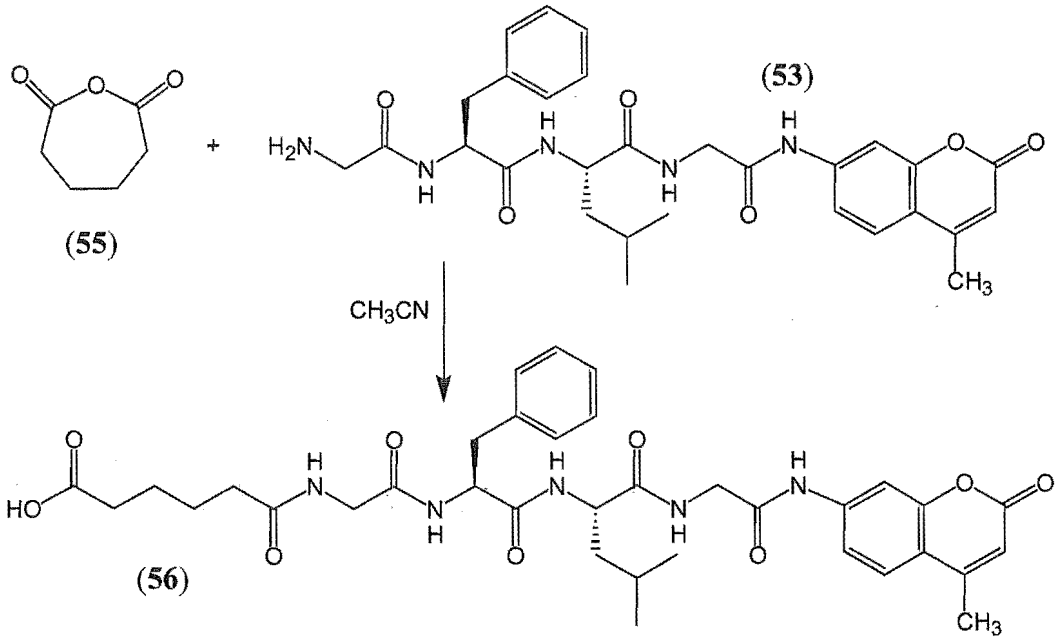


Figure 4.13 Adipic anhydride reaction with H_2N-Gly-Phe-Leu-Gly-coumarin (**53**)

4.3.5 Activation of Adipic-Gly-Phe-Leu-Gly-Coumarin

This reaction was the crucial step, as this was the point at which the previous strategies had failed with the 'ring closing' mechanism. Initial reactions using DSC (**54**) as the activating agent showed very little formation of the expected product (**Figure 4.14**) by ESIMS analysis of the crude reaction mixture. Therefore, DCC (1 equiv.) and *N*-hydroxysuccinimide (NHS, 1 equiv.) were added to **56** in DMF and left stirring at room temperature (**Figure 4.14**). An ESIMS sample analysis from the reaction mixture did show significant formation of a compound representing the correct mass for succinimidyl-adipic-Gly-Phe-Leu-Gly-coumarin (**57**). After 12 hours reaction time preparative HPLC (C18, 45% CH_3CN/H_2O (0.05% TFA)) was used to separate the starting material from the

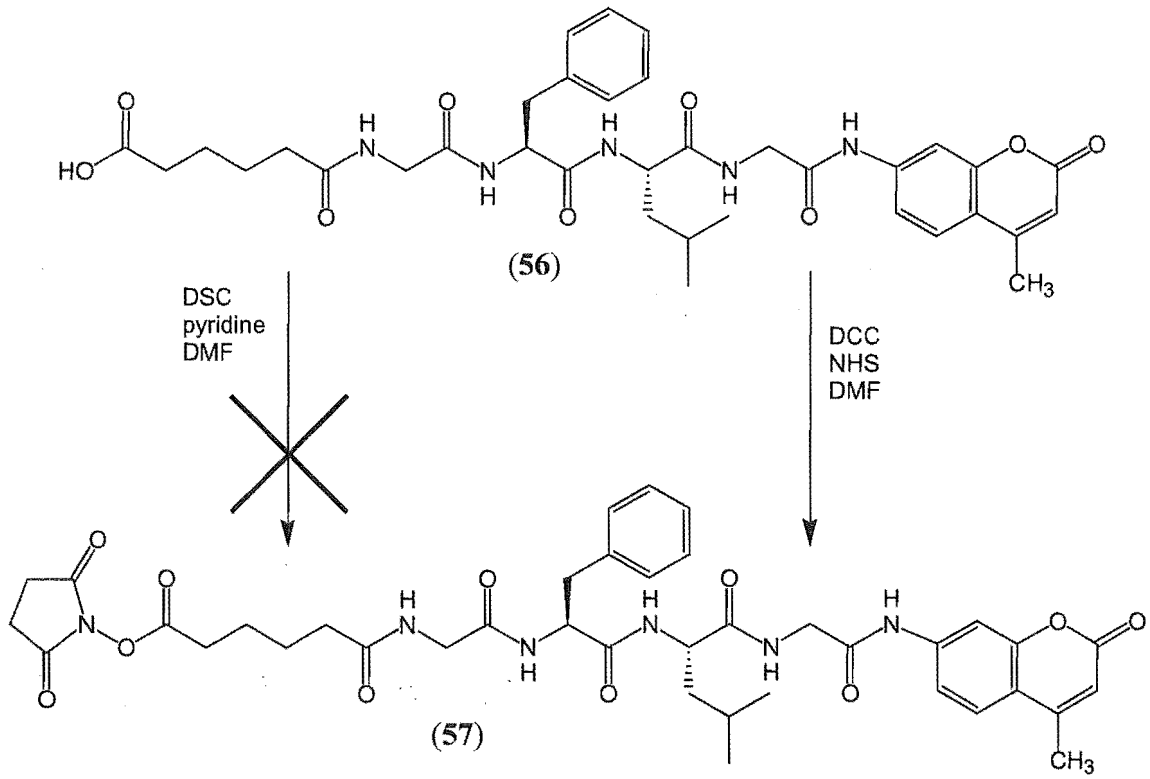


Figure 4.14 Activation of adipic-Gly-Phe-Leu-Gly-coumarin (**56**)

product (retention time 14.1 min), both of which showed the correct masses for their respective peaks. There was not enough material for any meaningful NMR analysis and previous experience with the succinimidyl esters had shown that they were best used immediately, before hydrolysis or other side reactions could destroy the ester. Therefore, the final reactions with CV-N (**15**) were carried out immediately.

4.3.6 Reaction of CV-N with 57

Based on the experience gained with reaction of CV-N (**15**) with the fluorescent dye derivatives in section **4.2**, the best approach for generating these conjugates was to load as much 'coumarin' onto each CV-N (**15**) as possible to ensure the best fluorescent response for later testing. Therefore, dual reactions of 25 equiv. of **57** with CV-N (**15**) were reacted in $NaHCO_3$ buffer with two reaction times of 15 and 60 min (**Figure 4.15**). After each reaction time was completed the samples were chromatographed on a G25 column eluting water. Detection at 276 nm ensured collection of the protein material separated from the lower molecular weight components. The protein containing fractions were then freeze dried.

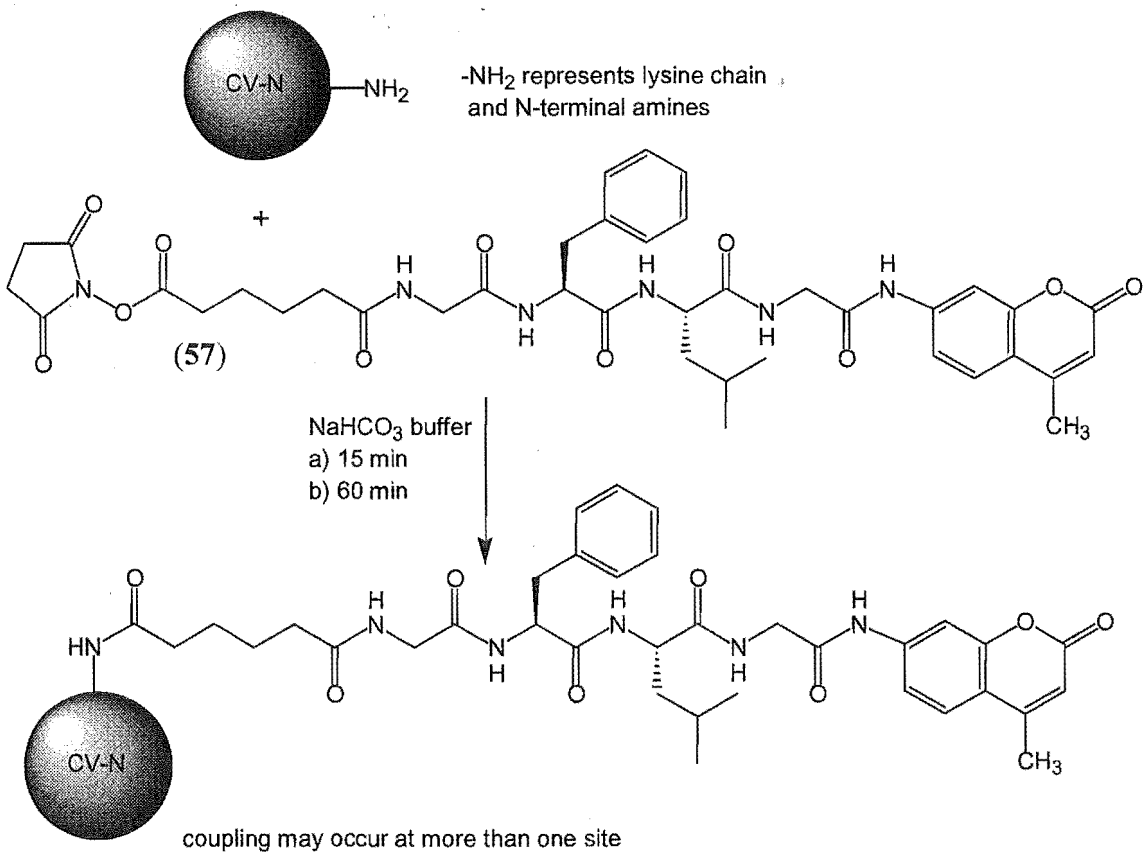


Figure 4.15 Reaction of CV-N (**15**) and **57**

4.3.6.1 Analysis of 15 minute Reaction

The product from the 15 min reaction (**58**) was dissolved in water and analysed by LCMS with injection onto a reverse phase C3 column running a gradient from 20-50% CH_3CN/H_2O (0.5% formic acid) over 15 min. Elution of the native CV-N (**15**) peak at 9.3 min was followed by the singly-conjugated CV-N (11.1 min) and the doubly-conjugated CV-N (12.7 min). The specific ESIMS spectra at these times are shown in **Figure 4.16**. The combined, transformed data (from 9.1 to 12.7 min) is shown in **Figure 4.17** and represents the relative amounts present of the full mass of the protein conjugates derived from the individual spectra. The effectiveness of the LCMS analysis over the direct inject ESIMS analysis can be seen in the increased signal intensity of the peaks resulting in a flat baseline and more accurate results compared to **Figures 3.4** and **3.5**. The mass difference observed between series 'A', 'B' and 'C' in **Figure 4.16** corresponds to the added mass (~659 Da) of the adipic-Gly-Phe-Leu-Gly-coumarin conjugate as expected. The ratio of the three products as shown in **Figure 4.17** demonstrates only partial conjugation had taken place after only 15 min reaction time as the major component was still native CV-N (**15**).

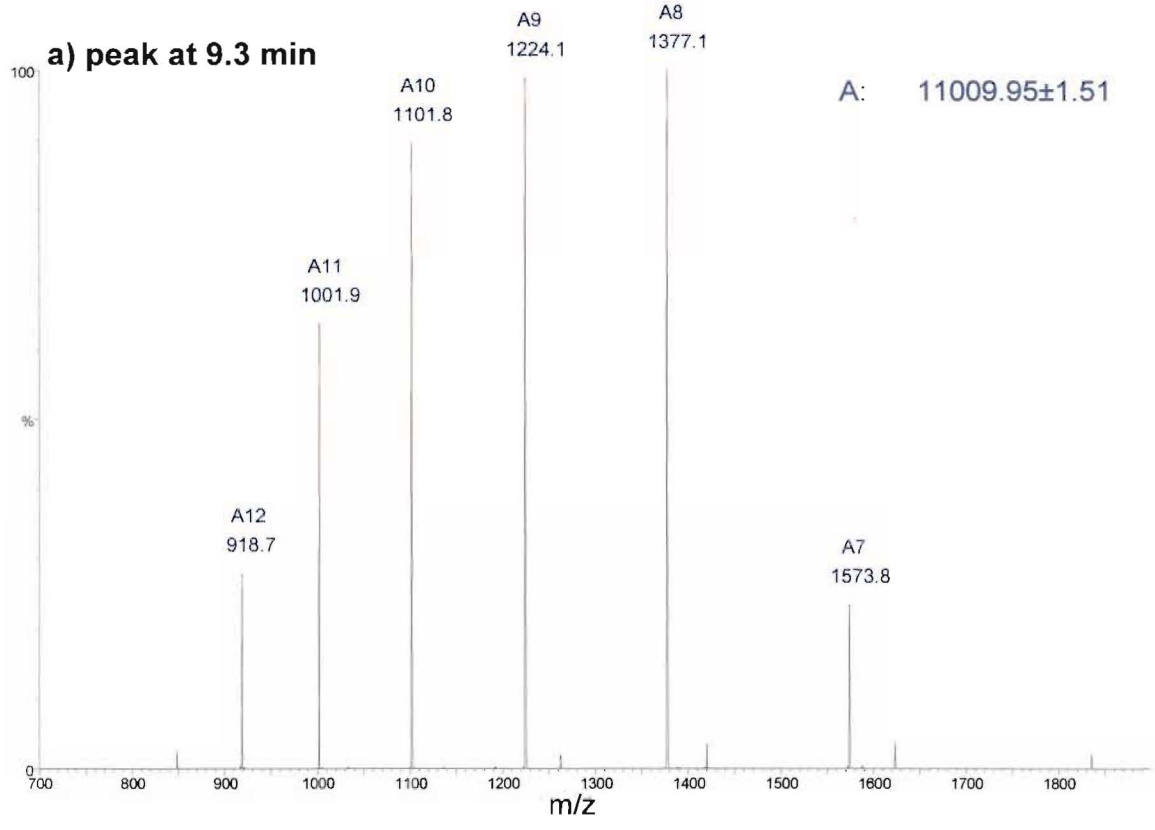


Figure 4.16a) LCMS spectrum of CV-N (**15**) plus **57**

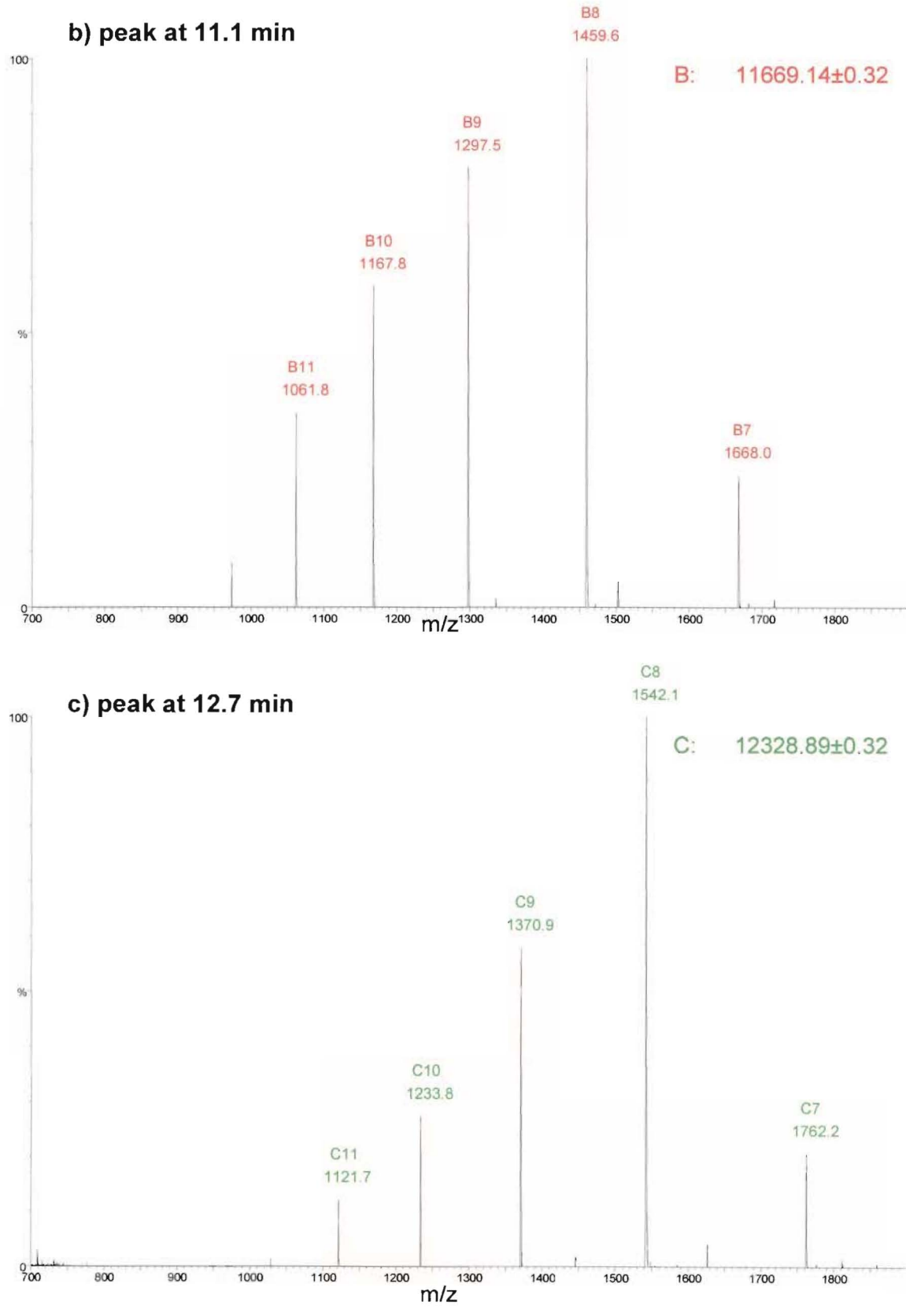


Figure 4.16b)-c) LCMS spectra of CV-N (**15**) plus 57

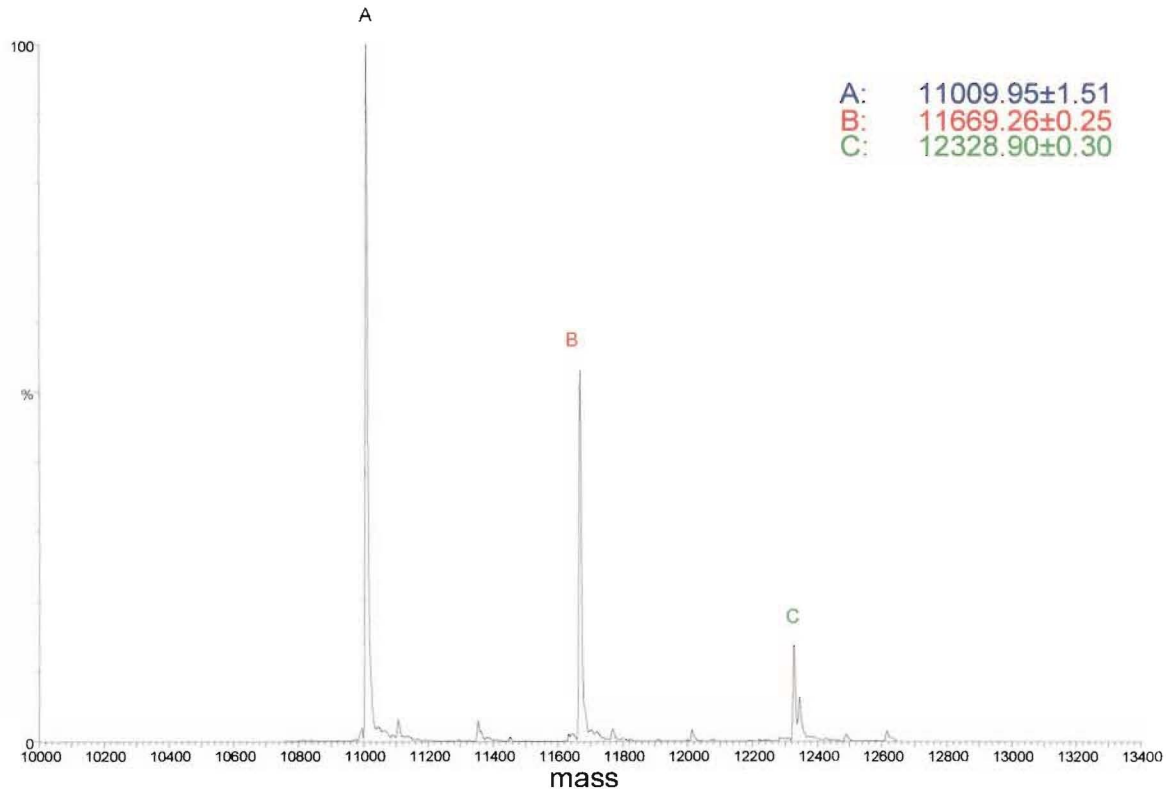


Figure 4.17 Transformed mass spectrum of CV-N (**15**) plus **57**

4.3.6.2 Analysis of 60 minute Reaction

The analysis of this sample (**59**) was performed as above. The individual components of the LCMS spectrum are shown in **Figure 4.18** with the combined, transformed data in **Figure 4.19**. The retention times for the peaks differ slightly from those of **Figure 4.16** as the samples were run on different days. The extended reaction time has allowed a greater degree of conjugation to the CV-N (**15**) molecule with up to four additions of the coumarin derivative. The transformed spectra (**Figure 4.19**) established the major component as the singly-conjugated CV-N with steadily decreasing amounts of the more conjugated species. There was still a substantial amount of native CV-N (**15**) present which ideally would be completely reacted in order to provide fully conjugated protein. In order to increase the rate of conjugation, the reaction time could be increased although hydrolysis of the succinimidyl ester could become significant. The second alternative would be to increase the ratio of succinimidyl ester (**57**) to CV-N (**15**).

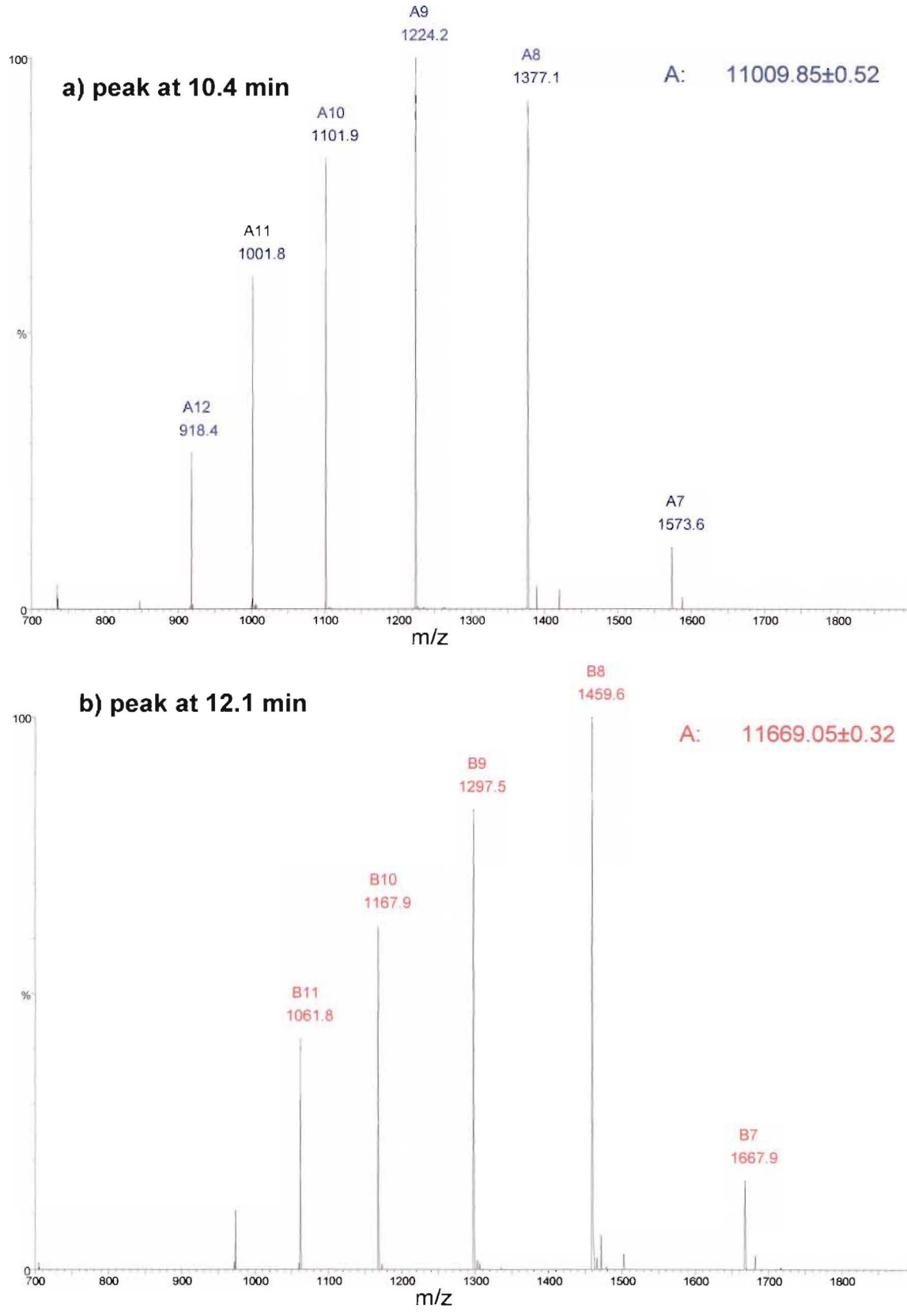


Figure 4.18a)-b) LCMS spectra of CV-N (**15**) plus **57**

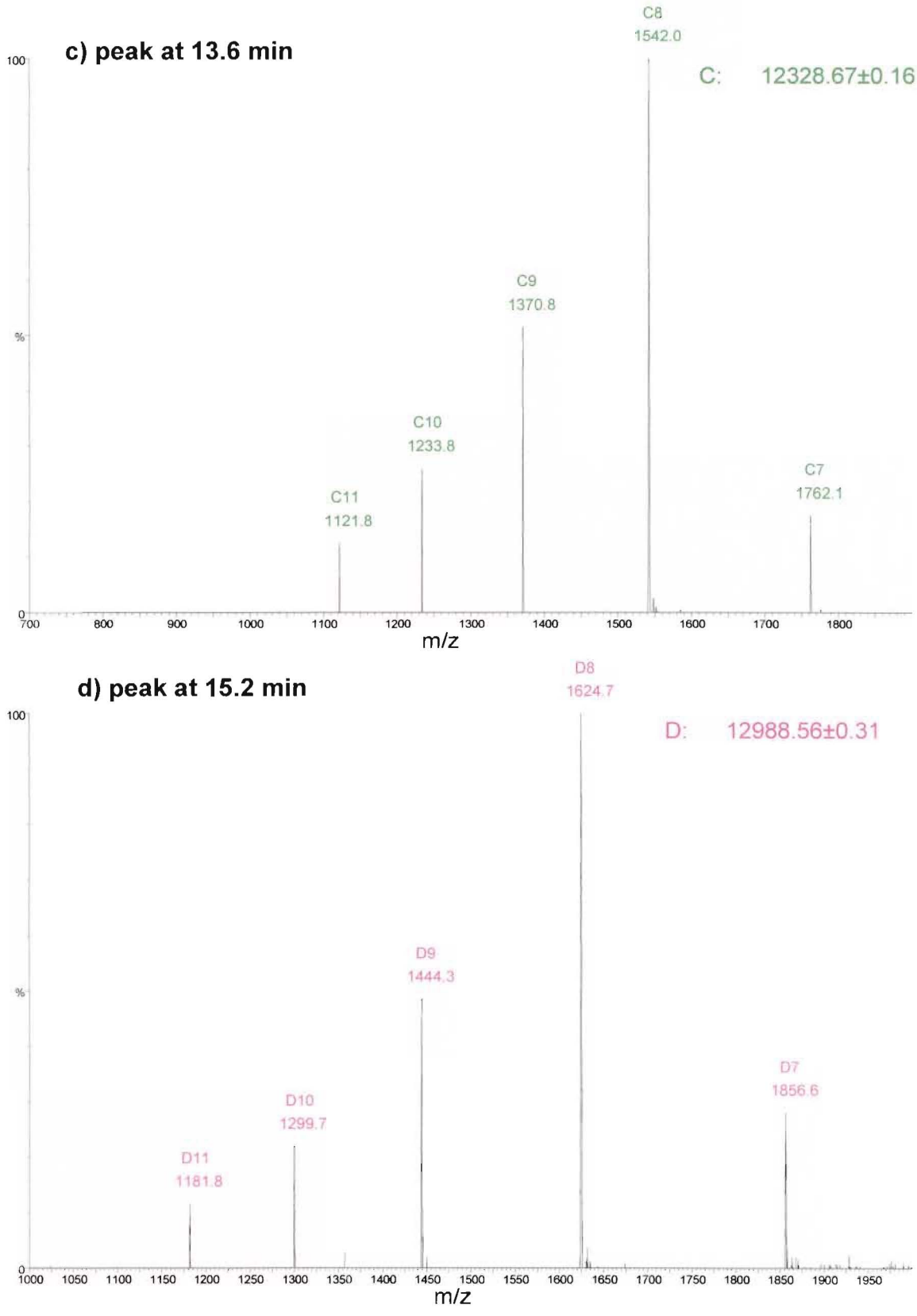


Figure 4.18c)-d) LCMS spectra of CV-N (**15**) plus **57**

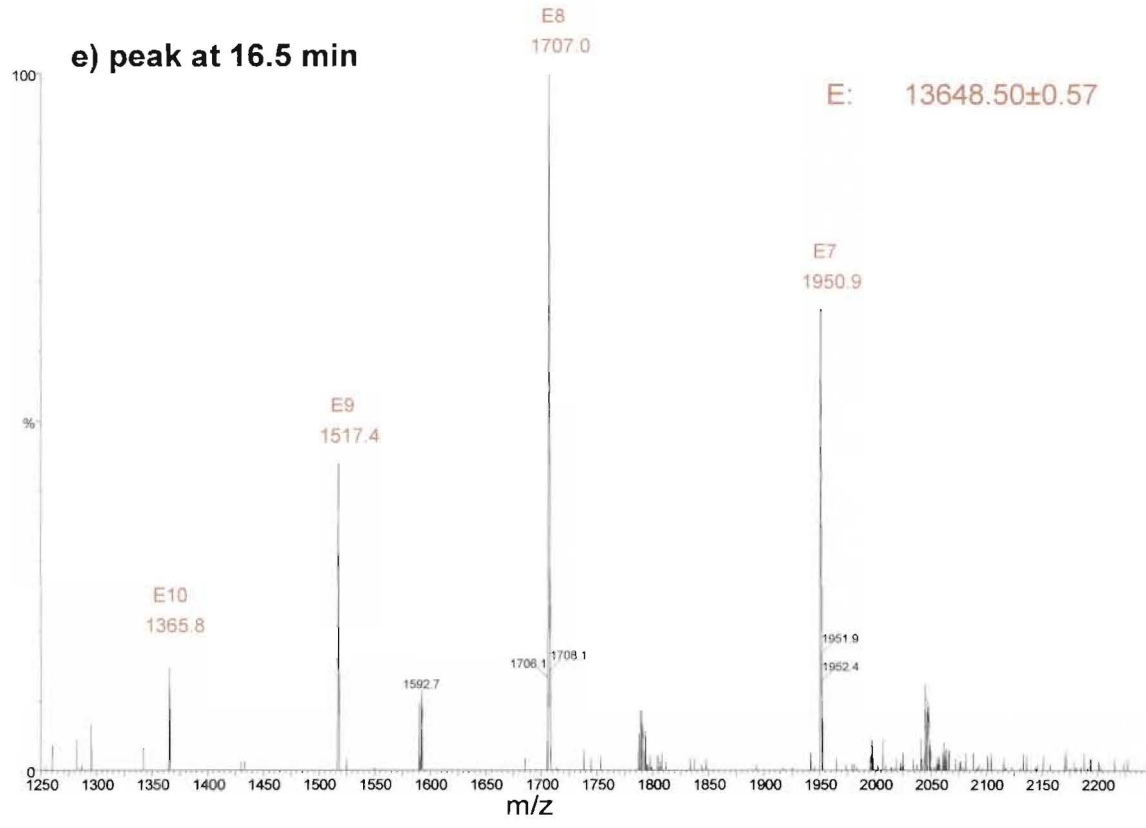


Figure 4.18e) LCMS spectrum of CV-N (**15**) plus **57**

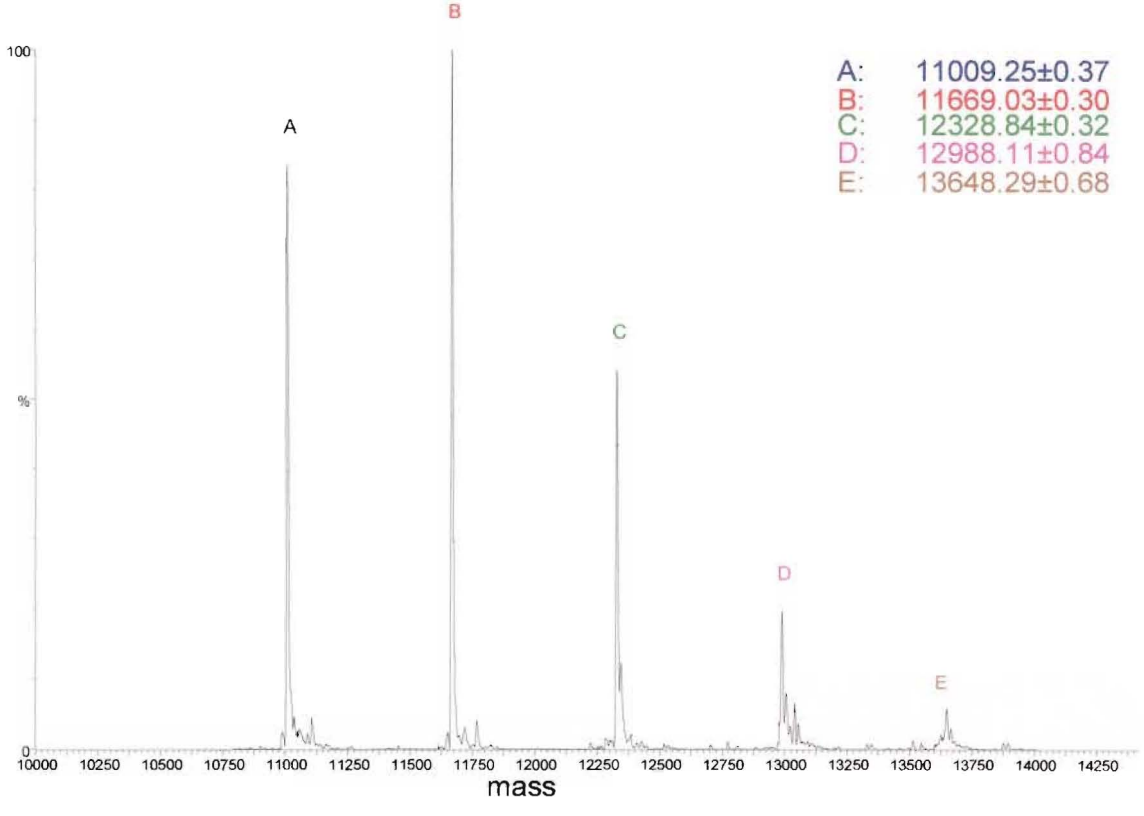


Figure 4.19 Transformed mass spectrum of CV-N (**15**) plus **57**

Both the CV-N conjugated coumarin samples underwent further testing (section **6.4**). The XTT assay results from these tests indicated, that like the CV-N-dye conjugates, these coumarin derivatives also bound gp120 in a comparable manner to the native CV-N (**15**). This confirmation of the binding meant that the final toxin conjugate synthesis could proceed as planned.

4.4 Succinimidyl-Biolinker-Toxin Constructs

The synthetic strategy developed in the previous section was now applied to construction of the identical conjugates with the replacement of the coumarin moiety with the halichondrin toxin. The harsh conditions required to attach the coumarin amine to the *C*-terminal of the biolinker were not needed for this conjugate as the aliphatic nature of the homohalichondrin B amine (**28**) would provide a greater degree of nucleophilicity.

4.4.1 Synthesis of Fmoc-Gly-Phe-Leu-Gly-Succinimidyl Ester

The carboxyl end of the Fmoc-biolinker needed to be activated in order for the amine of the halichondrin to be attached. The succinimidyl ester was chosen as a suitable activating functional group as the displacement reaction takes place under mild conditions with formation of an unreactive side-product (NHS). Reactions to activate the carboxyl were first attempted with DSC (**54**) but much better results were obtained with the following synthetic procedure shown as the first reaction in **Figure 4.20**.

The Fmoc-Gly-Phe-Leu-Gly-OH (**49**, 1.12 mg) was dissolved in fresh, dry, THF with DCC (2 equiv.) and NHS (2 equiv.). The reaction vessel was flushed with argon, sealed and left stirring. Analysis of samples at various time intervals by HPLC (C18, 60% CH_3CN/H_2O (0.05% TFA)) and ESIMS indicated that complete conversion had been achieved in six hours. Attempted purification of the Fmoc-Gly-Phe-Leu-Gly-succinimidyl ester (**60**) by either normal phase column chromatography or preparative reverse phase HPLC led to hydrolysis of the ester to return the starting material. Therefore, as the activated ester appeared quite labile, the next reaction was undertaken *in situ* without prior purification.

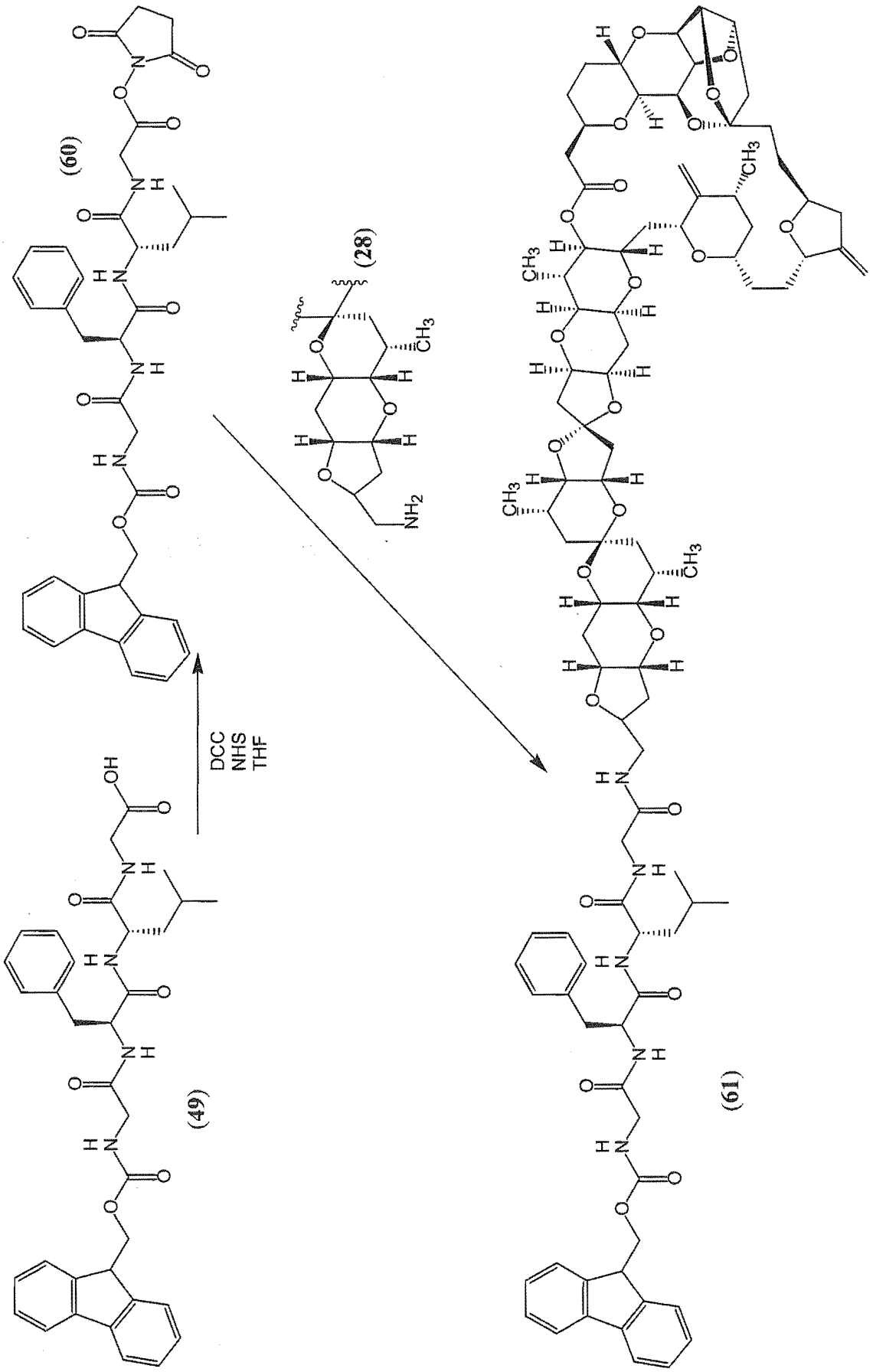


Figure 4.20 Synthesis of Fmoc-Gly-Phe-Leu-Gly-norhomohalichondrin B (**61**)

4.4.2 Synthesis of Fmoc-Gly-Phe-Leu-Gly-Norhomohalichondrin B

To the solution containing the activated ester **60**, was added norhomohalichondrin B amine (**28**, 0.5 equiv) in THF. The excess of the ester was to ensure complete reaction of all toxin amine which was more precious than the more readily synthesised peptide biolinker. The reaction progress was followed by ESIMS analysis which gradually showed the disappearance of the halichondrin amine and formation of the Fmoc-Gly-Phe-Leu-Gly-norhomohalichondrin B product (**61**) shown in **Figure 4.20**. The reaction was quite slow, taking 24 hours before all **28** had reacted. At this stage, HPLC (C18, 85% CH_3CN/H_2O) purification separated the product which was collected and dried. ESIMS analysis of the purified sample gave an accurate mass corresponding to the molecular formula of the product ($(MH_2)^{2+}$, $C_{94}H_{123}N_5O_{23}$) within 4.7 ppm error.

The ESIMS spectra of **61** is shown in **Figure 4.21**, the top spectrum representing the full mass range with the doubly charged region in the spectrum below. ESIMS analysis allows formation of adducts of sodium and potassium where the samples have been stored in glass. This leads to the many ion species represented in **Figure 4.21**. The three lowest doubly charged states shown (817.9, 826.9 and 835.9 Da) result from consecutive loss of H_2O in the mass spectrometer as this leads to conjugation and charge stabilisation, hence is favourable. The isotopic pattern for each species was due to the elemental composition of **61**, with the M+1 isotope peak dominating the pattern observed for the ions generated during ESIMS analysis. This is because as the number of carbons in a compound increases the natural ^{13}C abundance (~1%) becomes significant.

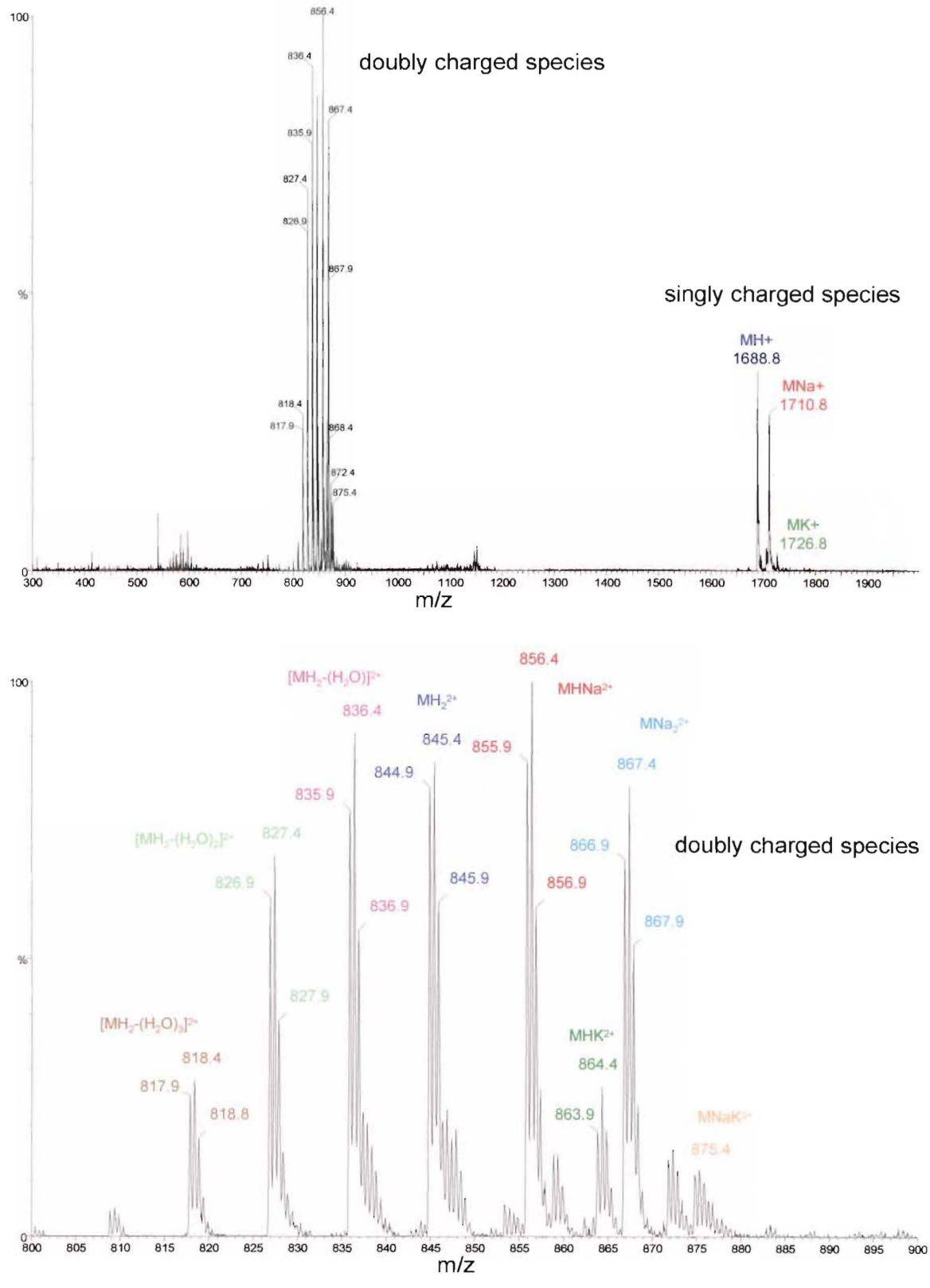


Figure 4.21 ESIMS spectrum of Fmoc-Gly-Phe-Leu-Gly-norhomohalichondrin B (**61**)

4.4.3 Fmoc Removal from 61

The next three synthetic steps are shown in **Figure 4.22** and represent conversion of the Fmoc-Gly-Phe-Leu-Gly-norhomohalichondrin B (**61**) to the activated succinimidyl ester after lengthening of the chain with adipic anhydride. The standard conditions for removal of an Fmoc group are a solution of 20% piperidine but in an effort to keep reaction conditions mild for the halichondrin moiety, it was discovered that a 2% piperidine/CH_3CN solution was sufficient. The Fmoc protected biolinker-halichondrin **61** was dissolved in 2% piperidine/CH_3CN and left at room temperature for 30 min. The solvent and piperidine were removed and the sample dried. An ESIMS analysis of the sample showed the correct mass for the H_2N-Gly-Phe-Leu-Gly-halichondrin species (**62**) with no sign of the starting material. The Fmoc-piperidine side product present was inert with respect to the next reaction and so the mixture was further reacted without purification.

4.4.4 Adipic Anhydride Reaction with 62

Fresh adipic anhydride was prepared (as per section **4.3.4.1**) and 1.1 equiv added to the H_2N-Gly-Phe-Leu-Gly-halichondrin (**62**) in dry CH_3CN. The sample was flushed with argon, sealed and left at room temperature. Consecutive ESIMS sample analyses (**Figure 4.23**) after 10, 20, 60, 120 min, and 26 hours, showed the slow conversion of amine to carboxyl (**63**) *via* the ring-opening adipic anhydride reaction. The $MHNa^{2+}$ and MNa_2^{2+} adducts of the starting material were the predominant ESI ion species initially seen, while the product formed primarily the MNa_2^{2+} ion. These data indicated that full conversion had taken place.

Purification of **63** was needed at this stage to remove unreacted adipic anhydride and/or adipic acid that was present in the product. The product was dissolved in minimal MeOH (~200 uL) and 50 uL aliquots chromatographed on a C18 cartridge (100 mg) with the empty space above the packing filled with H_2O (~ 5 column volumes) to dilute out the MeOH when added. Each aliquot was loaded on to the column in this way and then

washed with water. The cartridge was then washed with MeOH to elute **63** which was dried under vacuum. An accurate mass ESIMS analysis confirmed the molecular formula of the product ($C_{85}H_{119}N_5O_{24}Na_2$, 1.93 ppm error).

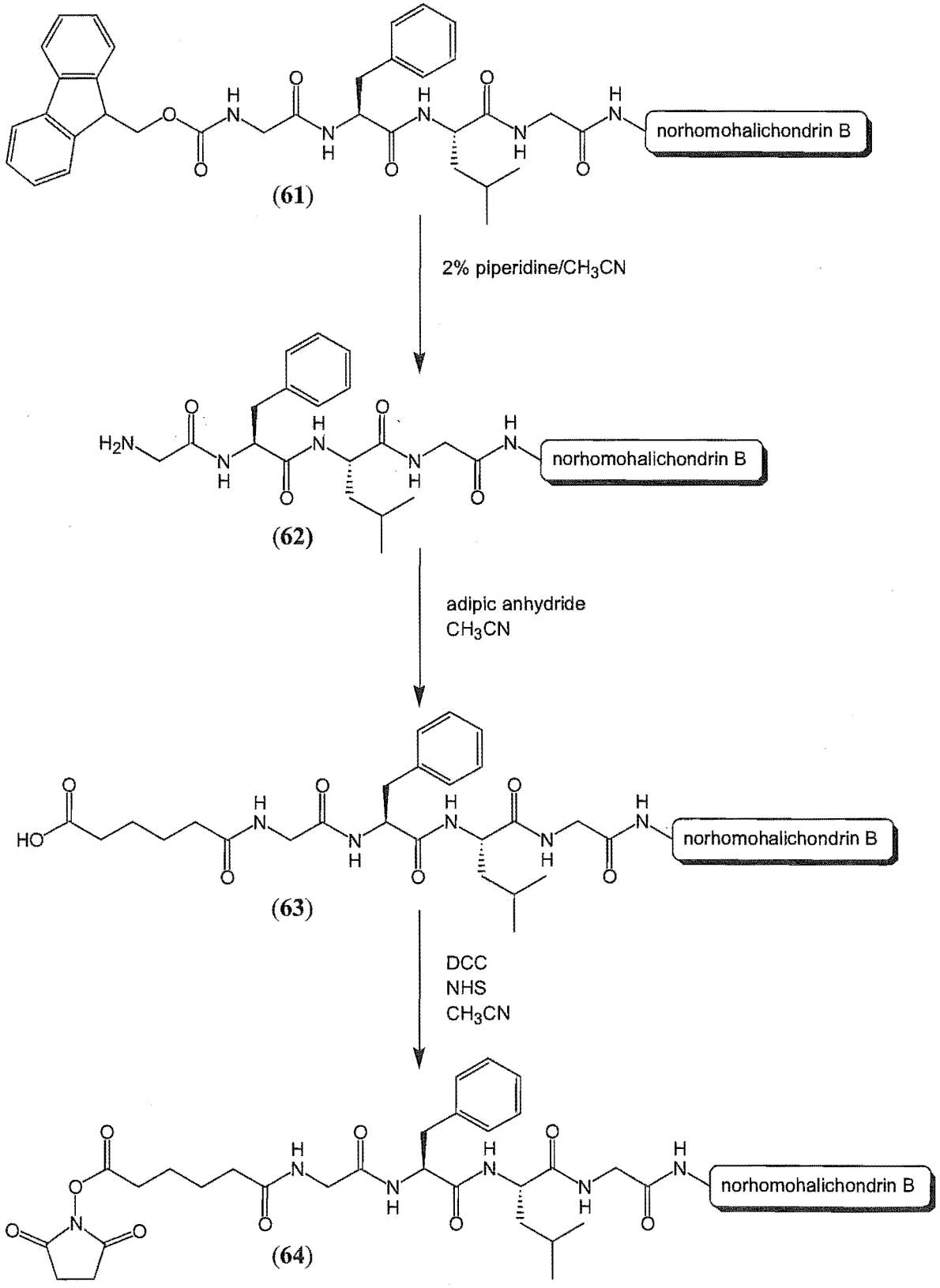


Figure 4.22 Synthetic scheme for conversion of **61** to **64**

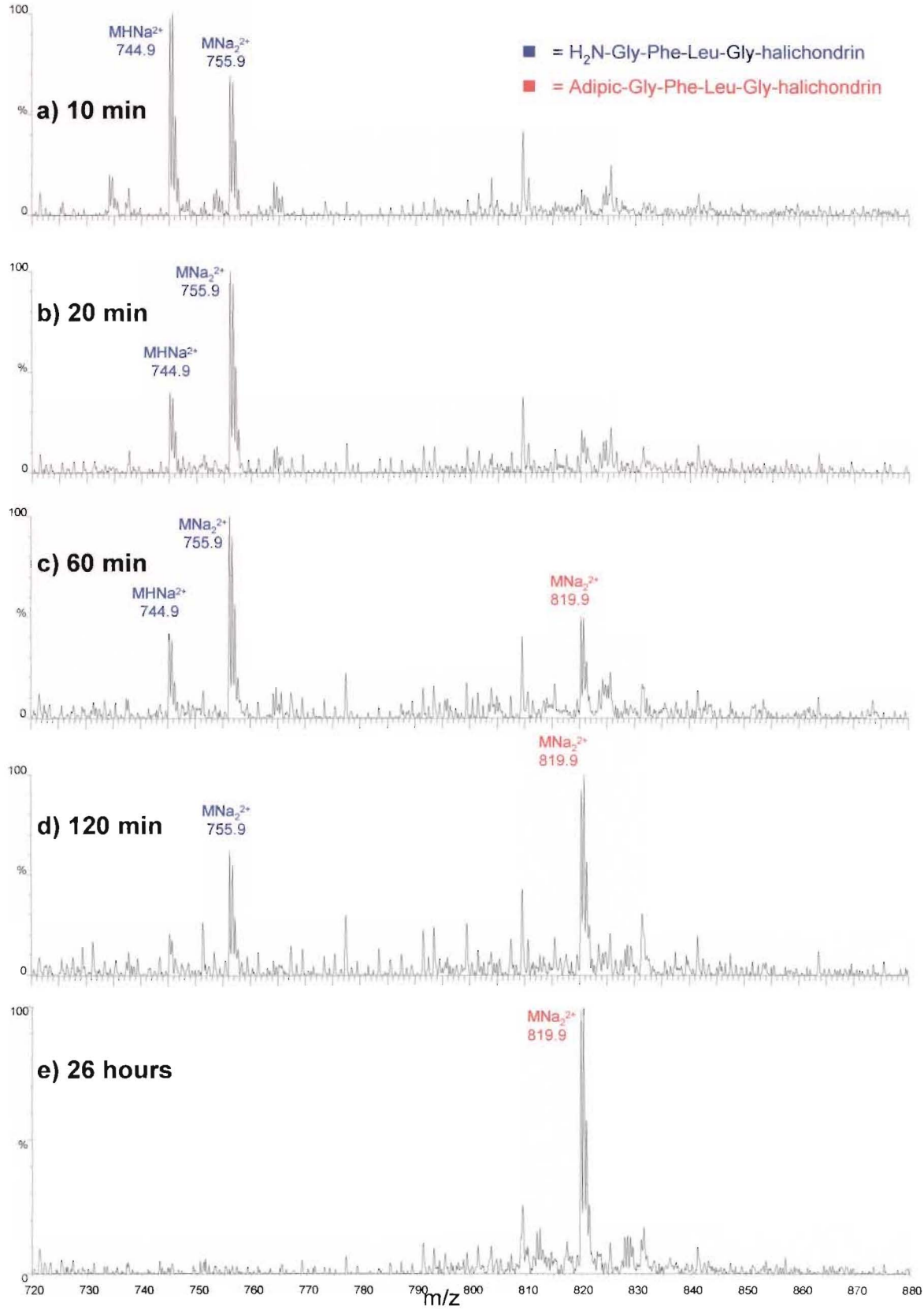


Figure 4.23 ESIMS time resolved formation of adipic-Gly-Phe-Leu-Gly-norhomohalichondrin B (**61**)

4.4.5 Activation of Adipic-Gly-Phe-Leu-Gly-Norhomohalichondrin B

The final reaction before coupling of the conjugate with CV-N (**15**) was to activate the free carboxyl of **63** with a succinimidyl ester. To the carboxyl derivative (**63**, 0.62 mg) was added DCC and NHS (5 equiv. each) in dry CH_3CN. An ESIMS spectrum after 48 hours showed complete conversion of the starting material to the succinimidyl ester (**64**). This product was purified by reverse phase C18 HPLC (65% CH_3CN/H_2O) (retention time of 6.5 min). After drying, the mass and molecular formula was confirmed by ESIMS analysis.

4.4.6 Reaction of 64 with CV-N

In the final reaction (**Figure 4.24**), CV-N (**15**, 50 μg) was dissolved in 150 μL of $NaHCO_3$ buffer (0.1 M) and **64** (20 equiv.) added in 50 μL of CH_3CN. The reaction was left at room temperature for 1 hour and an aliquot analysed by LCMS. The gradient was 20-60% CH_3CN/H_2O (0.5% formic acid) over 20 min. Formation of the product was observed although only one adipic-biolinker-toxin conjugate had been added. The separation of the compounds was excellent. A retention time of 9.5 min was observed for CV-N (**15**) with elution of the conjugate at 13.1 min. A further LCMS analysis after 15 hours showed a slight increase in the amount of the conjugated product present, but this was still far below the amount of native CV-N (**15**) present. Therefore, a further 20 equiv. of **64** was added to improve the yield of the reaction. LCMS analysis 15 hours after the addition of the extra toxin construct did show increased formation of the singly-conjugated product. A comparison of the ion series for each of the two components of the reaction mixture (**Figure 4.25**) shows the change from a mass of 11,009 Da (CV-N, **15**) to that of the product, 1,576 Da higher at 12,585 Da. The advantage of having such a large non-polar biolinker-toxin molecule attached was the dramatic shift in HPLC retention time of the constructs allowing separation of the ion series, as well as ready physical separation of the compounds themselves. Therefore, preparative HPLC on a C3 column with the same gradient profile as for LCMS, (except for 0.001% formic instead of 0.5%) separated the product from unreacted CV-N (**15**).

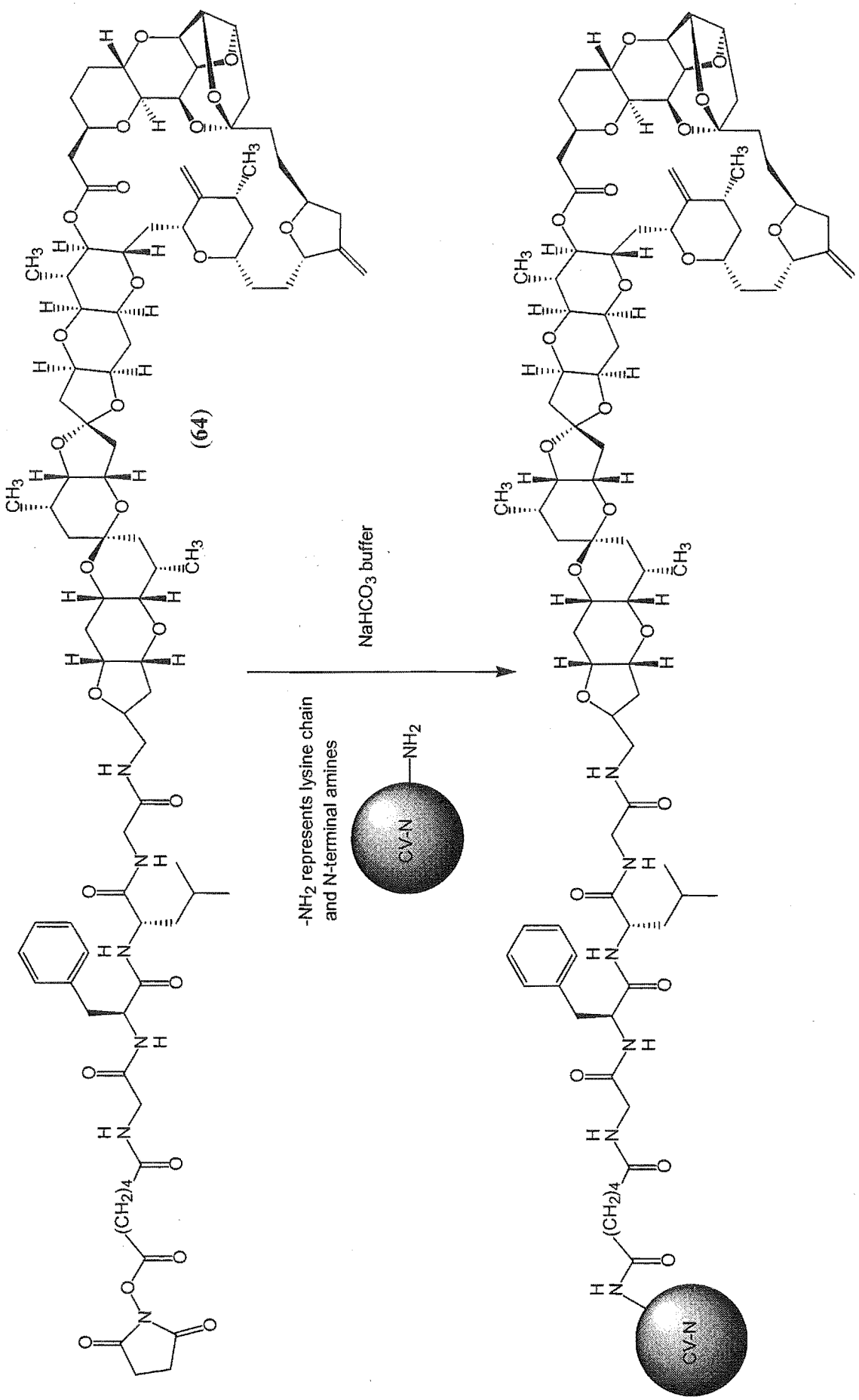


Figure 4.24 Synthesis of CV-N-adipic-Gly-Phe-Leu-Gly-norhomohalichondrin B

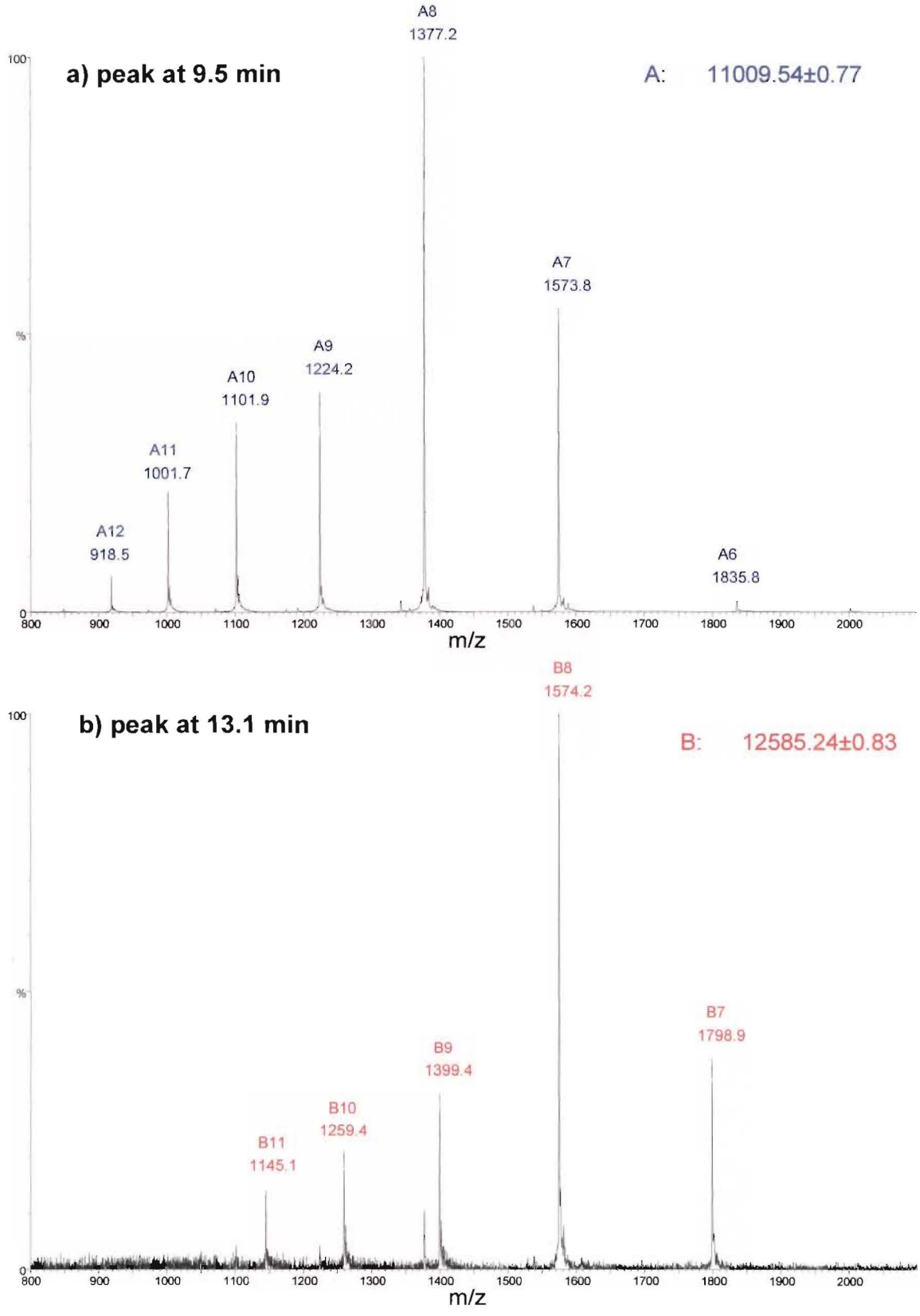


Figure 4.25 LCMS spectra of CV-N (**15**) plus **64**

4.4.7 Conclusions from the Toxin Conjugate Synthesis

The strategy ultimately used for the successful synthesis of the CV-N-toxin conjugate involved more steps than was initially envisaged. There was little choice but to lengthen the chain using adipic anhydride as any kind of intramolecular 'ring closing' as observed for the equivalent coumarin synthesis would have been undesirable. The enforced chain extension can in fact be seen as advantageous with the increased distance between the large toxin and the surface of the CV-N providing flexibility in allowing CV-N to still bind to gp120. Also, the reactions chosen were effective in that they did not interfere with the halichondrin structure by forming unwanted side products or by degrading the toxin. This has been an important consideration in determining viable strategies. Separation of the conjugated CV-N from the native protein was afforded by the relatively non-polar nature of the adipic-biolinker-halichondrin conjugate and although unforeseen was a welcome outcome of attaching such a large molecule.

Chapter 5

MALEIMIDE METHODOLOGY

5.1 Introduction

The complementary approach to the use of succinimidyl esters in developing CV-N-conjugates, was to utilise maleimide functionalities (**Figures 2.4** and **2.5**) and react them with thiol groups either present in CV-N-Cys, or synthetically added *via* iminothiolane (**27**). The production of CV-N-Cys (section **3.3**) and the development of CV-N (**15**) lysine amine conversion to thiols *via* iminothiolane (**27**) (section **3.4**) have been previously discussed. The other side of these strategies is the linking of the maleimido, biolinker, and toxin moieties together in an efficient synthetic scheme. The strategy to achieve this was the basis of this chapter.

Before full-scale maleimide derivative development could take place the reactivity of the maleimide group towards thiols was tested. The viability of the maleimido-thiol reaction was carried out first with a lysozyme model. The thiolated-lysozyme (formed in section **3.4.1**) was reacted with a simple maleimide derivative and the product analysed by ESIMS. The maleimide chosen for the equivalent reaction with thiolated-CV-N was a fluorescent compound (F-150, Molecular Probes Inc.) incorporating a maleimido-activating group and the same fluorescein dye moiety utilised in section **4.2**.

As with the development of the succinimidyl methodology, a test compound was first synthesised using the coumarin moiety in place of the toxin derivative. A method was adapted from the successful succinimidyl strategy to produce a maleimido-biolinker-coumarin adduct. The product formed was then reacted with thiolated-CV-N to produce CV-N constructs containing a dye attached through a biolinker.

Finally, the synthesis of the toxin equivalent was undertaken with the norhomohalichondrin B amine (**28**) and reaction with both CV-N-Cys and thiolated-CV-N carried through.

5.2 Thiolated Proteins with Maleimides

5.2.1 Thiolated-Lysozyme Reaction with a Maleimide

Maleimidobutyric acid (**65**) was chosen to react with the model protein, lysozyme, after conversion of the lysozyme lysine side chain amines to thiols using iminothiolane (**27**). The reaction mixture synthesised in section **3.4.1** contained multiple thiol groups and the reaction of these with a maleimide functionality was important to ensure that this reaction would be viable for CV-N (**15**). The reaction is outlined in **Figure 5.1**.

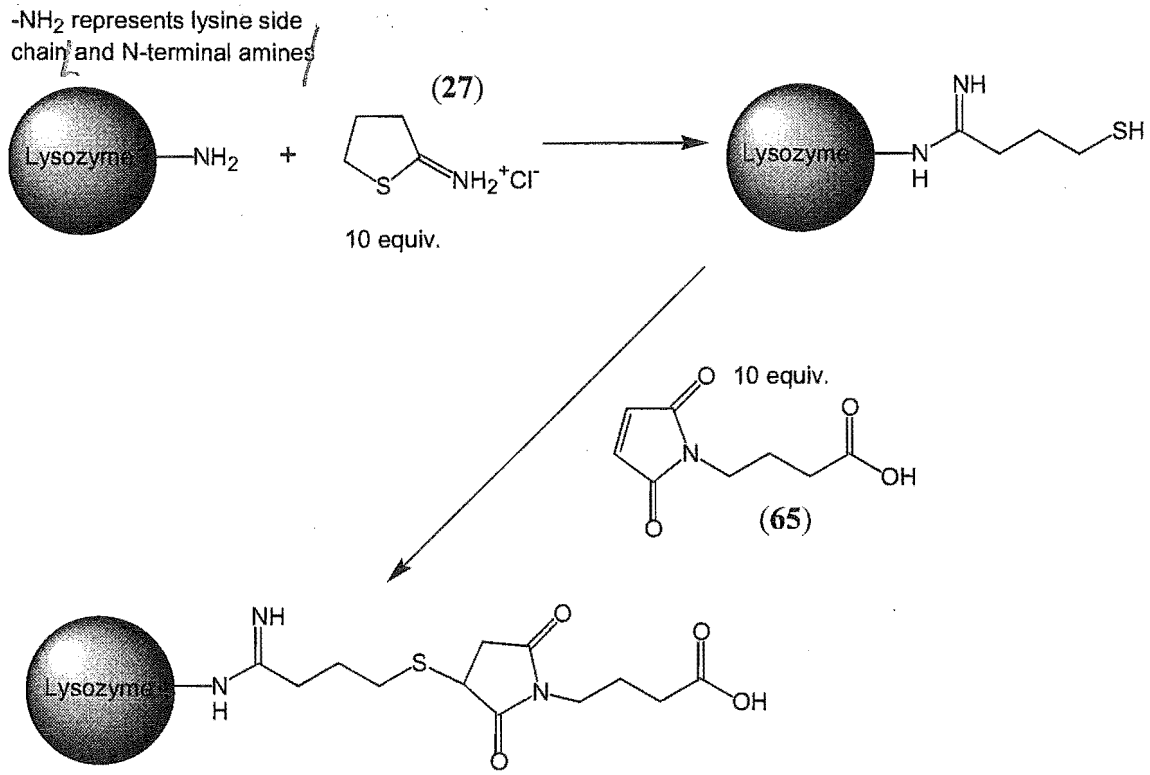


Figure 5.1 Lysozyme-thiolated model plus maleimido butyric acid (**65**)

Maleimidobutyric acid (**65**, 10 equiv.) in reaction buffer (50 μL, 0.1 M sodium borate, 0.001 M EDTA, 0.15 M NaCl, pH 8.0) was added to the thiolated-lysozyme mixture and the reaction left for 5 min. In contrast to the relative ease with which the thiolated-lysozyme ionised in the ESI mass spectrometer, after reaction with maleimidobutyric acid (**65**) the ion signal was markedly decreased. Addition of neat formic acid to the sample when injected assisted ionisation of the protein adducts to some degree and a transform of the spectrum is shown in **Figure 5.2**. The baseline is poor, but the consistent mass gap of 284 Da between components 'A', 'B', 'D', and 'E' corresponds to sequential addition of a maleimidobutyric acid group to a thiol on the lysozyme. Component 'C' corresponds to a lysozyme molecule having two thiols, only one of which has reacted with a maleimide.

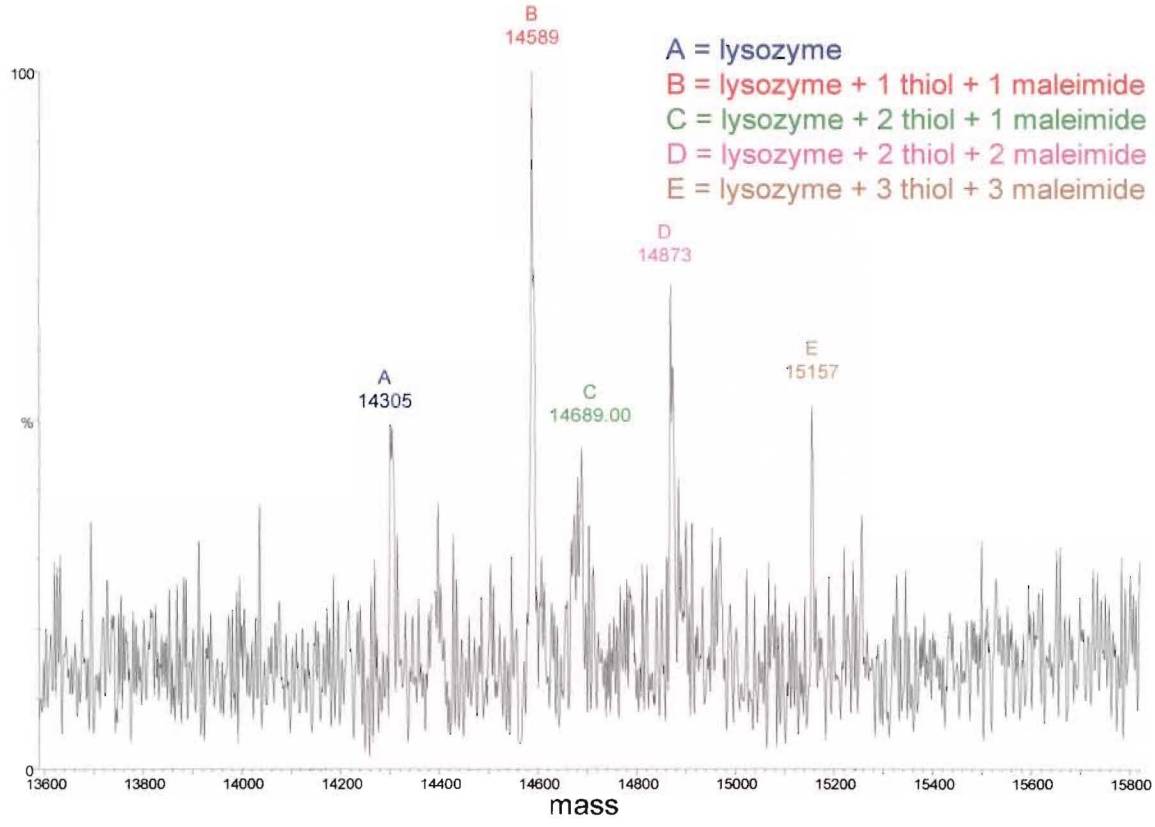


Figure 5.2 Transformed ESIMS analysis of lysozyme-thiolated plus maleimidobutyric acid (**65**)

The reaction between thiols and the maleimide group was fast with the majority of thiolated-lysozyme converted within 5 min. The relatively small size of the

maleimidobutyric acid undoubtedly played a role in this aspect of the kinetics of the reaction. The significant loss of ionisation of the protein conjugates after addition of the maleimide could be due to the side reaction of a hydrolysed iminothiolane with a maleimidobutyric acid molecule. This small side-product, very likely present in excess, could effectively quench the ionisation of the larger protein adducts. Therefore, this reaction points towards the use of LCMS and/or G25 column chromatography to more accurately survey the outcome of these type of reactions. The trade off is the length of time taken for the chromatography to take place for either method. However, the benefits of product separation and increased ionisation, hence a better signal to noise ratio, outweigh the 'speed' advantage of direct injection ESIMS analysis. This knowledge was applied in later reaction schemes.

5.2.2 Thiolated-CV-N Reaction with F-150

Having established that the model lysozyme system reacted successfully with the maleimidobutyric acid (**65**), the same reaction was attempted on thiolated-CV-N to more fully test the viability of the maleimide approach. The maleimide selected for reaction with the thiolated-CV-N mixture synthesised in section **3.4.2.1** was F-150 (**66**), a fluorescein derivative with a maleimide activating group. The synthetic scheme in **Figure 5.3** depicts the reaction with CV-N (**15**) after initial conversion of the lysine side chain amines to thiols (section **3.4.2.1**).

The F-150 dye (**66**, 70 equiv.) was added to the thiolated-CV-N (200 μg). After reaction for 1 hour the mixture was separated by G25 column chromatography (with monitoring at 276 nm) to remove the overwhelming excess of small molecules present. The protein-containing fraction was analysed by ESIMS. The resultant transformed spectrum is shown in **Figure 5.4**. There was still unreacted CV-N ('A') and thiolated-CV-N with 1 and 2 thiols added ('B' and 'C') although the majority of the thiolated-CV-N had reacted to form the F-150 fluorescent conjugates 'D', 'E' and 'F'. Although the number of equivalents of the F-150 dye was equal to that of the iminothiolane (**27**) eventually added to the CV-N

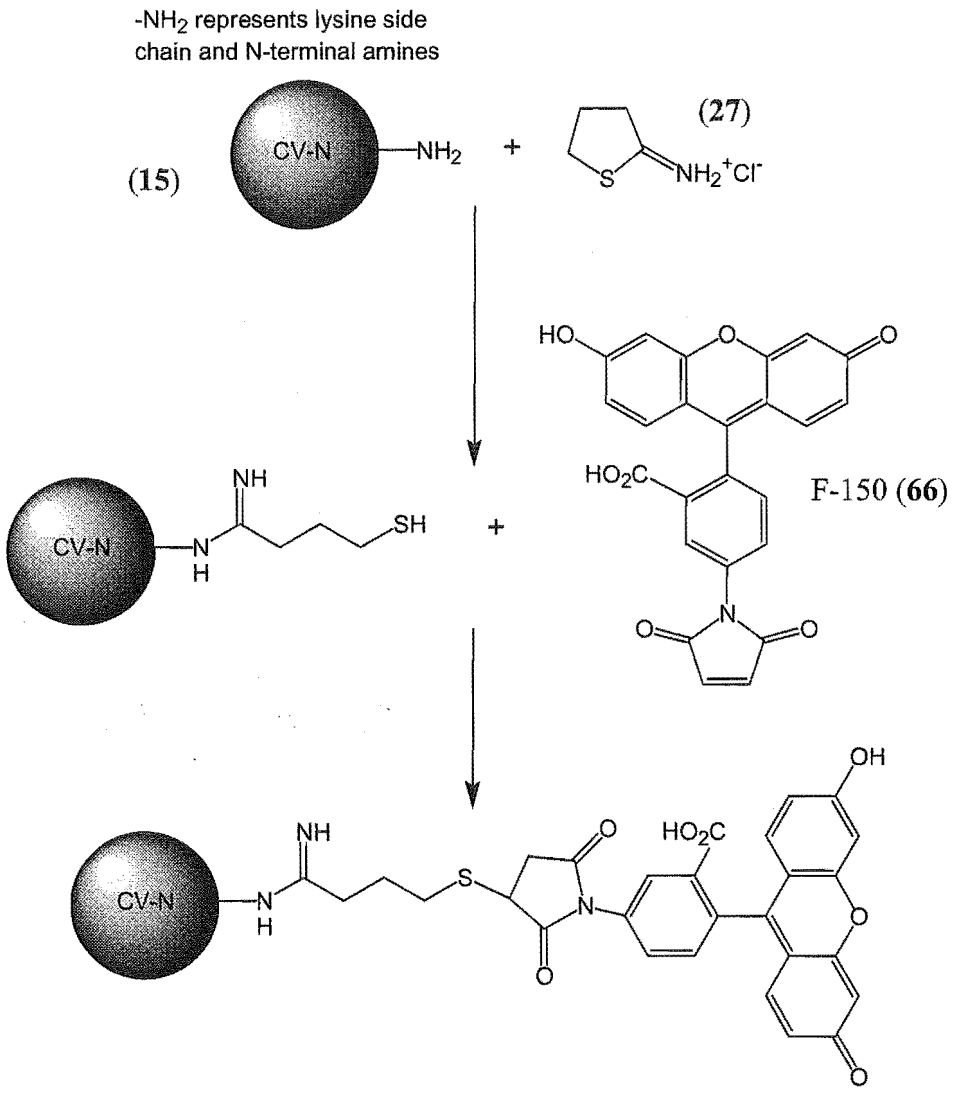


Figure 5.3 Reaction of thiolated-CV-N with F-150 (**66**)

(**15**), the presence of unreacted thiols was a cause for concern as potentially the toxin-loaded maleimide could be 'wasted' on reaction with hydrolysed iminothiolane, leading indirectly to incomplete conversion of the protein thiols. It was therefore concluded that the thiolated-protein should first be purified by G25 chromatography before reaction with the maleimide. This would remove the majority of the excess lower molecular weight compounds present and allow just the protein thiols to be reacted with the excess of the specific maleimide derivative.

The reactions of maleimides with thiolated-lysozyme, and then thiolated-CV-N, confirmed the validity of this approach for the attachment of conjugates to proteins. The results also

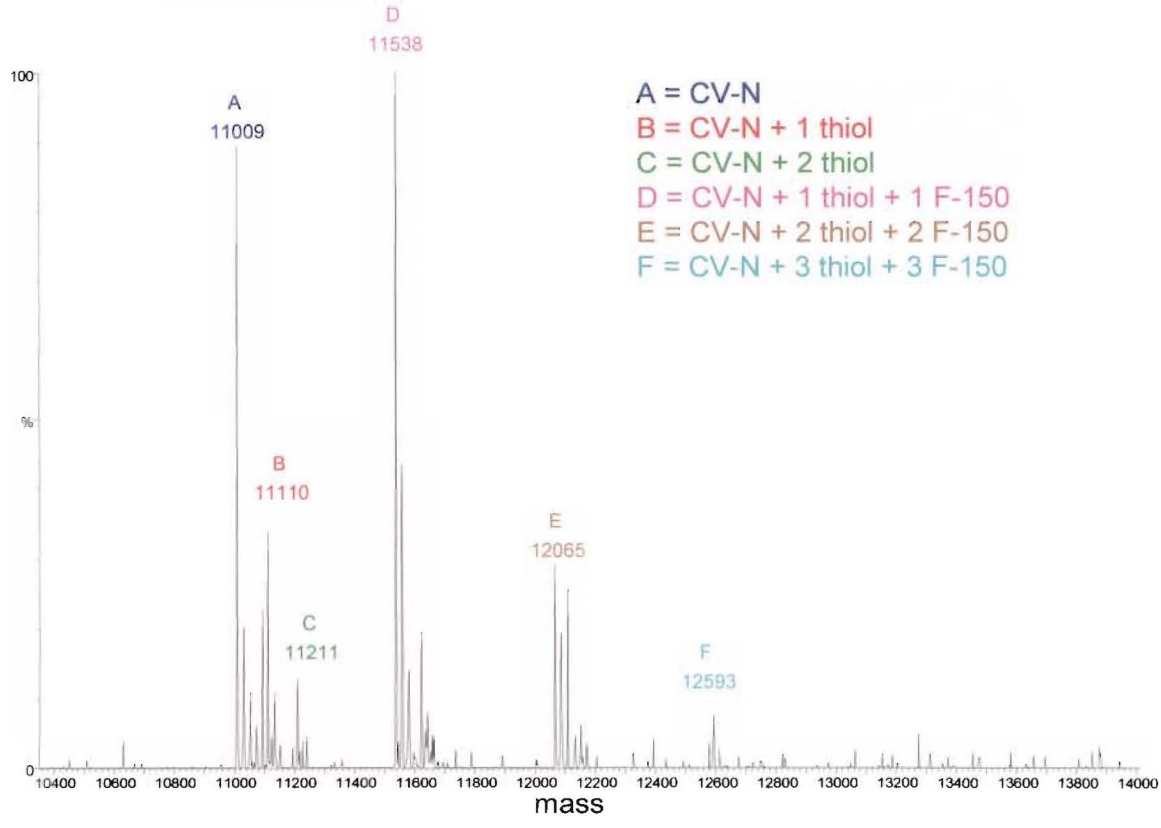


Figure 5.4 Transformed ESIMS analysis of thiolated-CV-N plus F-150 (**66**)

showed that direct injection ESIMS analysis is suitable only up to the point of addition of the maleimide to the thiolated-protein. This led to the more important consideration of utilising G25 chromatography for separating the excess hydrolysed iminothiolane (with a free thiol capability) from the thiolated-protein before addition of an overwhelming excess of maleimide to ensure complete reaction of the thiolated-protein. The products of the reaction would then be purified by G25 chromatography and/or analysed by LCMS.

5.3 Maleimido-Biolinker-Coumarin Constructs

Although the model reactions reinforced the applicability of the maleimide-thiol reaction to the design of CV-N conjugates, the synthesis of the biolinker-toxin had to be addressed. Unlike the succinimidyl esters, the activating group (maleimide) could be introduced early on in the synthesis as the introduction of the amino toxin would cause no reactivity problems. The succinimidyl ester strategy required early toxin addition and then subsequent modification to allow introduction of the activated ester. The advantage of utilising the maleimide was that the sensitive toxin would only be exposed to the reaction conditions to introduce it, and then followed immediately by reaction with thiolated-CV-N or CV-N-Cys.

The coumarin fluorescent dye was again utilised to provide a toxin substitute and the tetrapeptide biolinker was also retained. The maleimide derivative chosen was maleimidocaproic acid (**67**) which would give a $(CH_2)_5$ pendant chain separating the biolinker from the protein, comparable to the adipic acid moiety in the succinimidyl strategy. This gives a degree of flexibility to the whole conjugate and, importantly, keeps the toxin distal from the surface of the CV-N molecule. There was one obvious method for synthesis of the maleimido-Gly-Phe-Leu-Gly-coumarin (**68**), shown in **Figure 5.5**. The carboxyl of the maleimidocaproic acid (**67**) allowed a simple coupling reaction to take place to the free amine of the biolinker-coumarin construct.

5.3.1 Synthesis of Maleimido-Caproic-Gly-Phe-Leu-Gly-Coumarin

Synthesis of the maleimide derivative utilised the H_2N-Gly-Phe-Leu-Gly-coumarin (**53**, 2.0 mg) already synthesised in section **4.3.2.2**. To this was coupled maleimidocaproic acid (**67**, 1 equiv.) with the coupling reagent DCC (1.5 equiv.) in DCM (plus 3 drops of DMF for solubility). The reaction mixture was left stirring overnight. TLC analysis (20% MeOH/EtOAc) confirmed formation of a fluorescent compound less polar than the

starting materials. The solution was dried, taken up again in MeOH and purified by preparative HPLC (C18, 45% CH_3CN/H_2O (0.05% TFA), 324 nm) to yield **68**. HRESIMS analysis indicated a molecular formula of $C_{39}H_{46}N_6O_9$ as required for **68**.

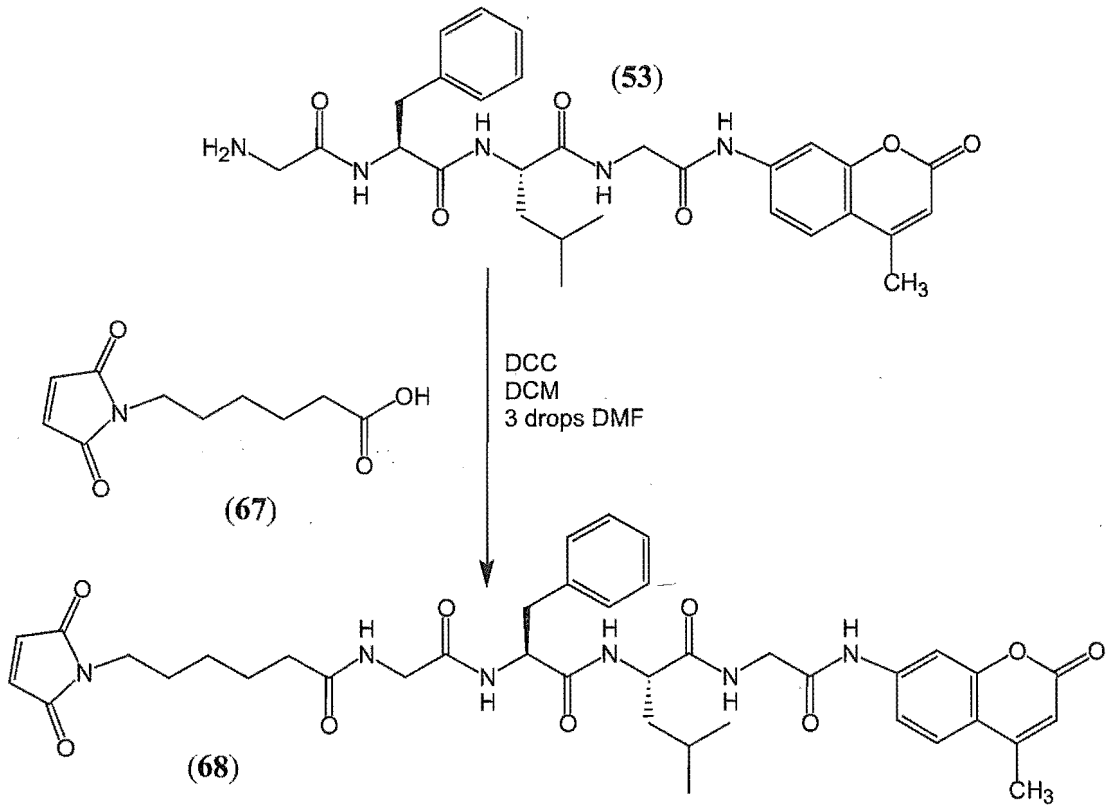


Figure 5.5 Synthesis of Maleimido-caproic-Gly-Phe-Leu-Gly-coumarin (**68**)

5.3.2 Thiolated-CV-N with 68

The thiolated-CV-N formed in section **3.4.2.2** had been purified by G25 column chromatography and was ready for reaction with a maleimide. The coumarin adduct (**68**, 20 equiv.) in DMF was added to the thiolated-CV-N (200 μg) in reaction buffer and left at room temperature. After 2 hours the solution was concentrated under a stream of N_2 and analysed by LCMS (20-50% CH_3CN/H_2O (0.5% formic) over 15 min). Native CV-N (**15**) that had not been thiolated was still present but sufficient quantities of the singly (11,853

Da) and doubly (12,696 Da) conjugated products could be identified with concomitant disappearance of the thiolated-CV-N species.

The synthesis and subsequent reaction with CV-N of the maleimide derivatives has demonstrated the simplicity of this strategy in comparison to that of the succinimidyl ester approach. To achieve the synthesis of the final activated species for reaction with CV-N (or derivatives) required much fewer reaction steps for the maleimides than for the succinimidyl esters. There is the additional step required for conversion of the CV-N (**15**) amines to thiols, but this was achieved simply and efficiently.

5.4 Maleimide-Biolinker-Toxin Constructs

The major difference in the synthetic strategy for the maleimido toxin derivatives compared to the equivalent coumarin constructs was that the starting material for this strategy needed to be maleimido-caproic-Gly-Phe-Leu-Gly-OH (**69**). The simple approach to this compound was to synthesise the entire structure by SPPS and simply cleave the peptide derivative from the resin. Attaching the toxin molecule to the maleimide-biolinker would then only require a single step of *C*-terminal activation and *in situ* displacement by reaction with norhomohalichondrin B amine (**28**). This strategy is shown in **Figure 5.6**.

5.4.1 Synthesis of Maleimido-Caproic-Gly-Phe-Leu-Gly-OH

The SPPS method used was identical to that for Fmoc-Gly-Phe-Leu-Gly-OH (**49**, section **4.3.1.2**) with removal of the terminal Fmoc group while still on the resin, and subsequent coupling of the maleimidocaproic acid (**67**) to the peptide chain. The removal of the peptide derivative with the TFA solution and subsequent purification was also identical. An ESIMS analysis of the crude product showed the correct mass, but HPLC analysis (C18, 60% MeOH/H_2O (0.05% TFA)) indicated the presence of several minor impurities. The product was purified by preparative HPLC using the same elution conditions. Analyses by HPLC and ESIMS were consistent with the structure. ^{1}H NMR chemical shifts were assigned from comparison to previously assigned biolinker derivatives. [It was noted that this product could also be used to prepare the coumarin derivative as an alternative to the synthesis in section **5.3.1**. This was indeed done but yields were lower than those for the synthesis detailed in section **5.3.1**.]

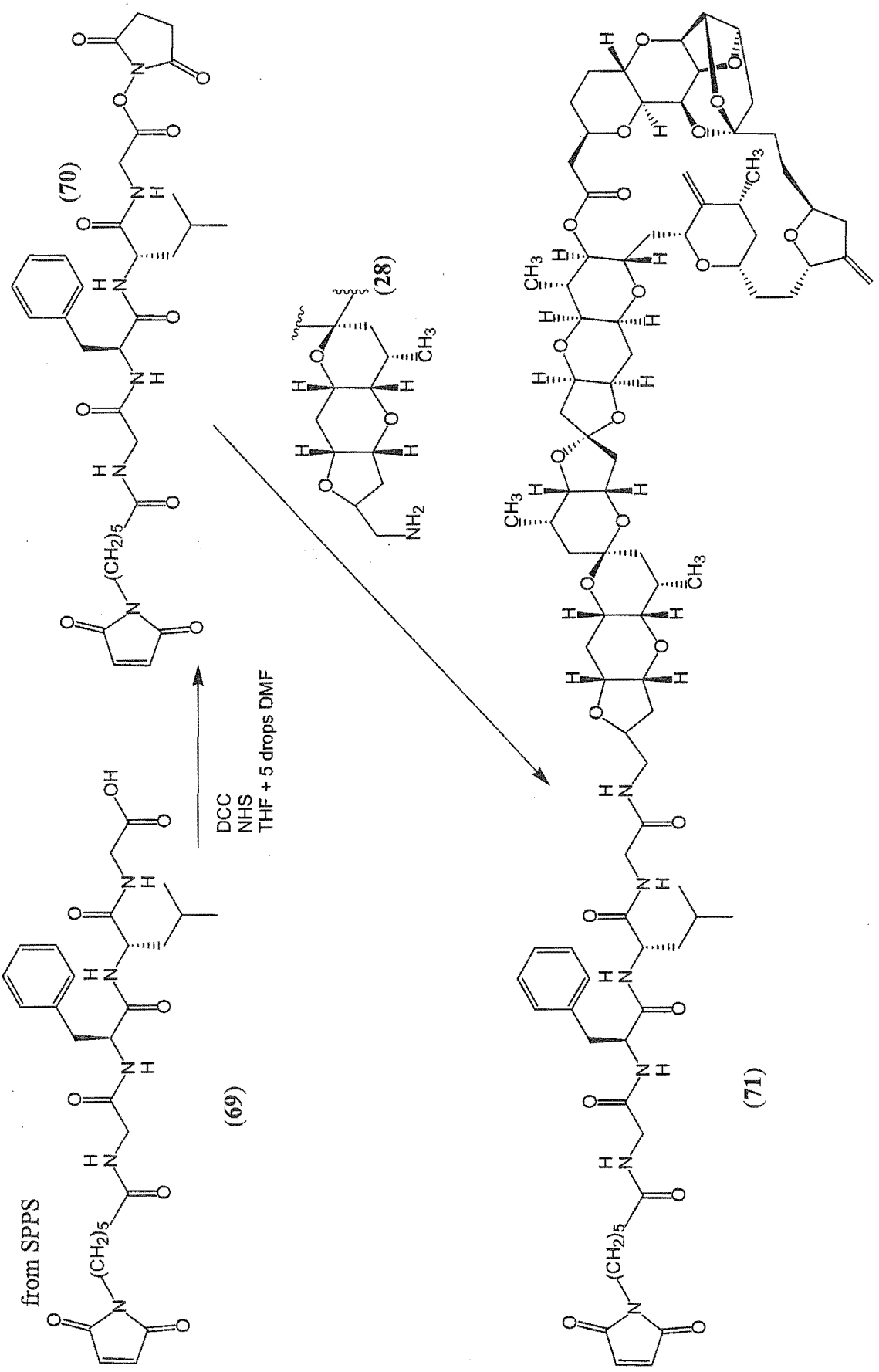


Figure 5.6 Synthesis of Maleimido-caproic-Gly-Phe-Leu-Gly-norhomohalichondrin B (**71**)

5.4.2 Synthesis of Maleimido-Caproic-Gly-Phe-Leu-Gly-Toxin

The activation of the carboxyl by formation of the succinimidyl ester and subsequent reaction with the toxin amine was carried out in the same reaction vessel to avoid hydrolysis of the succinimidyl ester. Maleimido-caproic-Gly-Phe-Leu-Gly-OH (**69**, 1.07 mg) was dissolved in dry THF with DCC and NHS (2 equiv. each). Several drops of DMF were added to aid dissolution of the maleimide and the solution stirred under argon. After 24 hours ESIMS analysis showed approximately 50% conversion to the succinimidyl ester (**70**) At this stage norhomohalichondrin B amine (**28**, 0.5 equiv.) was added in CH_3CN to the reaction mixture. The reaction was continuously monitored by ESIMS analysis until complete disappearance of the toxin amine peak and formation of the product (**71**) after 60 hours. A gradient HPLC elution profile (25-85% CH_3CN/H_2O) was used to purify the maleimido-caproic-Gly-Phe-Leu-Gly-norhomohalichondrin B product (**71**) from maleimido-caproic-biolinker-OH (**69**).

An accurate mass was obtained by HRESIMS analysis of the purified sample (**Figure 5.7**). As noted earlier for the Fmoc-Gly-Phe-Leu-Gly-norhomohalichondrin B (**Figure 4.21**), the compound had produced sodiated adducts resulting in a range of ions, both singly and doubly charged. To form the product required a reaction time of greater than 48 hours, longer than that required for the equivalent Fmoc-biolinker. These long reaction times were almost certainly due to the steric bulk of two large molecules having to orientate themselves correctly in solution for reaction to occur.

The major advantage of ESIMS analysis of reactions such as this is the small sample amount required to monitor the reaction progress. This must be balanced by the fact that not all compounds ionise identically so that the 'ion count' for each component of the spectra may not be indicative of the relative quantity of each species present.

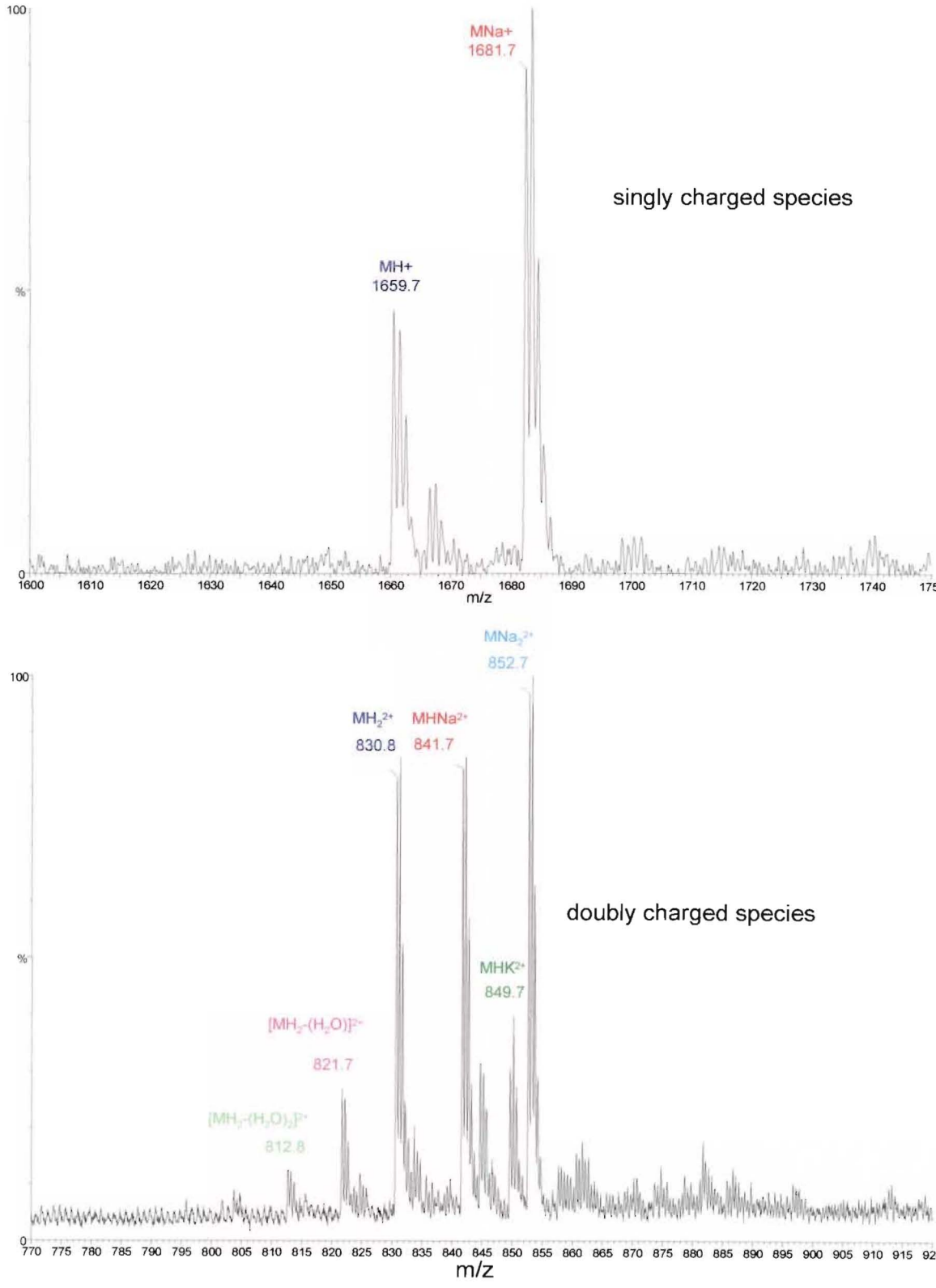


Figure 5.7 ESIMS spectra of maleimido-caproic-Gly-Phe-Leu-Gly-norhomohalichondrin B (**71**)

5.5 Maleimide-Toxin Reaction with CV-N Adducts

The utility of the maleimide-toxin derivative was that two methods were developed for its use. The first was reaction with CV-N-Cys where the free side chain of the extra cysteine residue provides a thiol for reaction at a single defined position. This would seem to be the most advantageous application of the maleimido-toxin derivative as the product formed would be a homogeneous compound. However, the alternative of pre-forming the thiols through conversion of the lysine amines to thiols would provide greater coverage of the CV-N with toxin molecules and give the potential for an increased cell killing ability. As was discovered in section **3.4.2.3** there was also the possibility of halting the thiolation reaction at a designated time when just a single thiolation adduct had been formed. The reaction for both CV-N substrates was identical as both were a maleimide-thiol interaction.

5.5.1 Maleimide-Toxin Reaction with CV-N-Cys

The recombinant production of the CV-N-Cys (section **3.3**) produced the correct compound but subsequently glutathione had coupled to the free thiol. The attempts at removal of the glutathione were partially successful in that the unwanted addition was easily removed but the refolding of the CV-N-Cys showed a loss of bioactivtiy suggesting that the refolding had not produced the correct tertiary structure of the CV-N. However, subsequent work at the LDDRD based on the reactions in section **3.3.3** did produce CV-N-Cys that was bioactive by ELISA studies (gp120 binding) and XTT assay (anti-HIV) results. This material was free of glutathione, but did contain several impurities which seemed to be addition of methyl groups (14 Da increase in mass) at indeterminate sites on the CV-N-Cys protein. The exact method of attachment of these 'methyl' groups was not determined. The fact that the material contained the free cysteine thiol and that the activity was equivalent to native CV-N (**15**) was enough to utilise this CV-N-Cys for reaction with the maleimide-toxin adducts.

Maleimido-caproic-Gly-Phe-Leu-Gly-norhomohalichondrin B (**71**,10 equiv.) was added to CV-N-Cys (**72**, 100 μg) in PBS buffer. The reaction was left overnight and LCMS analysis after 24 hours showed formation of about 15% product. The mixture was left for a further 24 hours before further LCMS analysis. The yield of product had not significantly altered, so the starting material and product were separated by reverse phase HPLC (C3 Zorbax, 20-60% CH_3CN/H_2O (0.001% formic) over 20 min). The fractions were dried and analysed separately by LCMS analyses to confirm the presence and identity of CV-N-Cys and the CV-N-Cys-biolinker-toxin construct.

The transformed LCMS spectra of the two compounds is shown in **Figure 5.8** with the **a)** spectrum representing the starting material CV-N-Cys ('A') and 'methyl' adducts ('B' and 'C') while the **b)** spectrum shows the product ('D') also with methyl adducts ('E' and 'F'). The fact that product formed meant that at least some of the CV-N-Cys had a free thiol on the cysteine as opposed to being S-methylated. The full potential of this strategy for forming CV-N-toxin conjugates will only be realised if CV-N-Cys can be formed free of impurities (such as glutathione or 'methyls'). In these circumstances a single homogeneous compound would be formed.

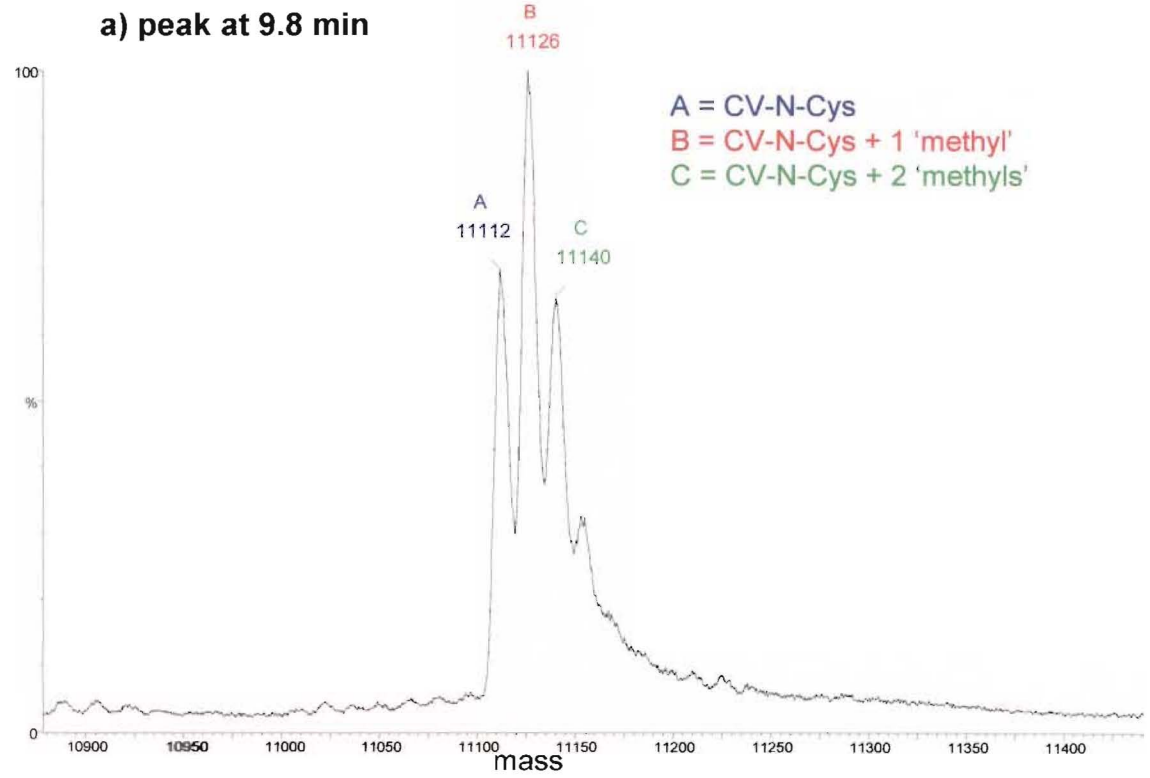


Figure 5.8a) Transformed LCMS spectrum of CV-N-Cys plus **71**

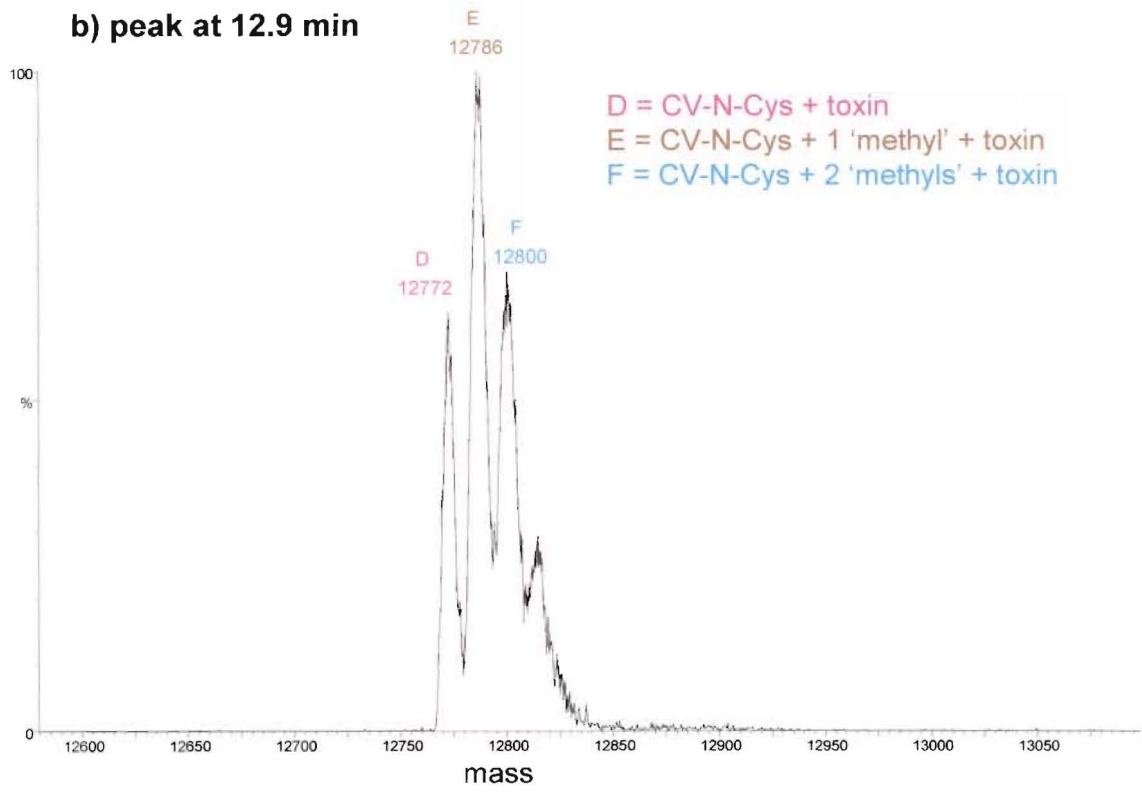


Figure 5.8b) Transformed LCMS spectrum of CV-N-Cys plus **71**

5.5.2 Maleimide-Toxin Reaction with Thiolated-CV-N

The thiolated-CV-N (section **3.4.2.3**) was primarily mono-thiolated as the reaction was halted after 30 min at which time this was the predominant species present. To this thiolated-CV-N (100 μg) in H_2O was added maleimido-caproic-Gly-Phe-Leu-Gly-norhomohalichondrin B (**71**) in DMF (25 μL). The reaction was left for 90 min and then analysed by LCMS utilising a 20-60% CH_3CN/H_2O (0.5% formic acid) gradient over 20 min. The starting material mixture of unreacted native CV-N (**15**) and thiolated-CV-N eluted at 9.7 min with the conjugate eluting at 13.1 min. Unreacted maleimide-toxin construct was also still present (eluting at 19.4 min) and so the reaction was left for a further 24 hours. LCMS analysis under the same conditions showed an increase in the product peak and the complete disappearance of the maleimide-toxin starting material. The LCMS spectrum of the starting material peak at 9.7 min was identical to that shown in **Figure 3.6**. A transformed spectrum of the product peak eluting at 13.1 min is shown in

Figure 5.9. The 'A' peak shows the addition of one iminothiolane group to CV-N with subsequent addition of a single maleimido-caproic-Gly-Phe-Leu-Gly-norhomohalichondrin B (**71**) group with a total mass of 12,771 Da. Therefore, the aim of stopping the initial thiolation reaction in order to obtain primarily a mono-thiol on CV-N and hence a single addition of toxin had been successful. HPLC under the same conditions (except replacement of TFA with 0.001% formic) separated the product from the remaining CV-N (**15**) and thiolated-CV-N.

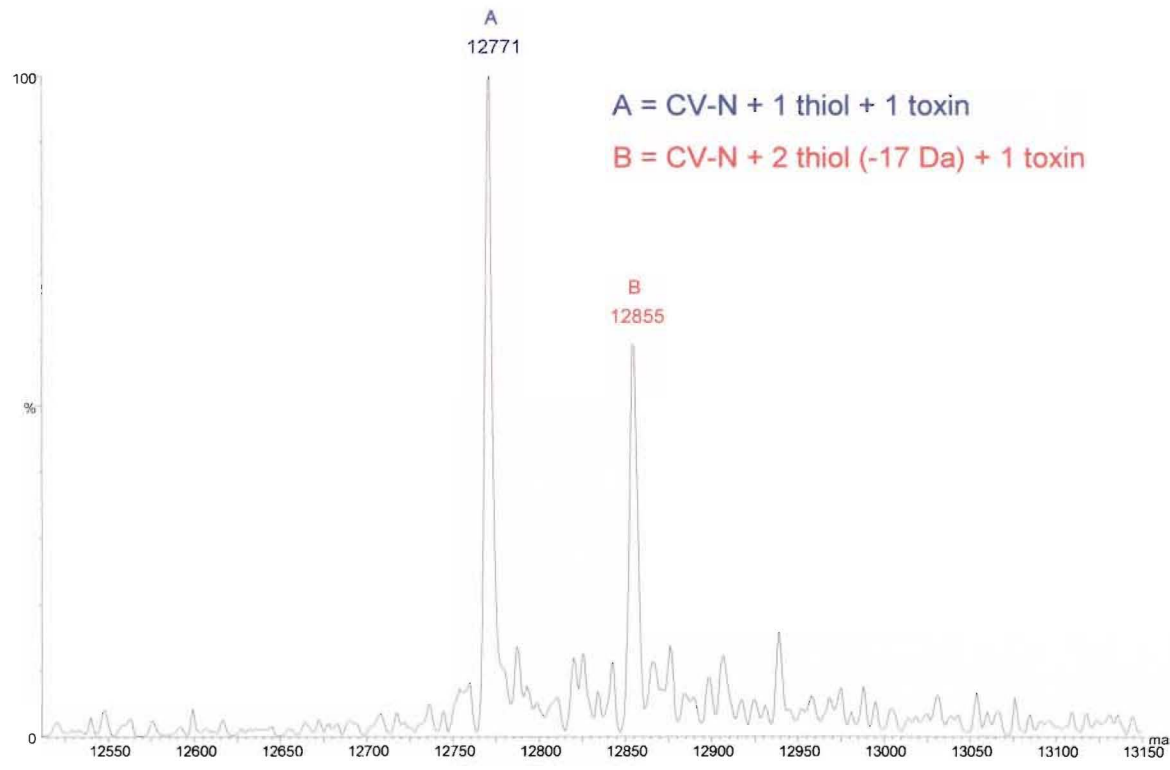


Figure 5.9 LCMS transformed spectrum of thiolated-CV-N plus **71**

The second ('B') peak in **Figure 5.9** had an interesting derivation. It was expected that there would be present a proportion of CV-N with two iminothiolane adducts and a single toxin derivative of total mass 12,882 Da. However, the peak seen in the spectrum is 17 Da less than this at 12,855 Da. Peaks 17 Da less than the mass for an iminothiolane addition had been seen previously at much lower levels in reactions of CV-N with iminothiolane (**27**). A potential structure for the side product is shown in **Figure 5.10**. This compound

matches the mass spectral data with a loss of 17 Da from the thiolated-CV-N although obviously there are other potential compounds that fit this criteria. It is suggested that this entity is formed by intramolecular reaction as indicated. It would be difficult to prove such

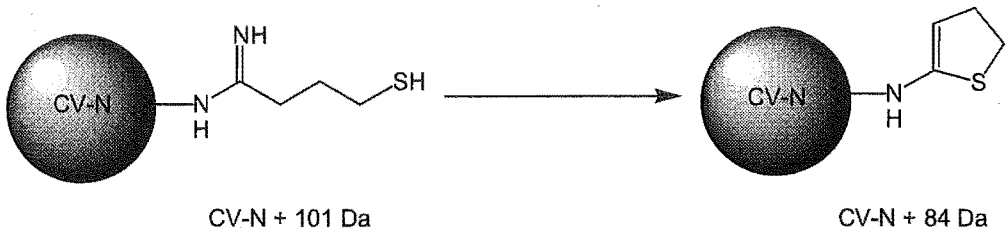


Figure 5.10 Postulated side product from thiolation of CV-N

a mechanism for the formation of the by-product and more work would need to be done to fully explain such a reaction. (This loss of 17 Da was only observed for thiolation reactions of CV-N (not lysozyme) and only occurred after successful formation of the correct thiolated product.) Evidence that supports this postulated structure is the absence of products containing two attached toxins. The presence of this non-reactive transformed lysine did not inhibit formation of the desired product and was therefore not further investigated.

It is crucial to note that by LCMS analysis, all the toxin adduct starting material had reacted after 24 hours confirming the inherent reactivity of the maleimide with thiols. The size and relatively non-polar nature of the biolinker-toxin construct again made separation of the product a simple task. A smaller toxin may not have contributed as much to the change in retention time and may not have led to complete separation by HPLC.

5.6 Conclusions from Maleimide Methodology

The early reactions with both lysozyme and CV-N with lower molecular weight maleimides demonstrated the speed and selectivity of the thiol-maleimide reaction. The synthesis of the maleimido-caproic-biolinker-coumarin conjugate was rendered effective through adaptation of the previously derived H_2N-biolinker-coumarin compound. The extension of the chain through the caproic linker mimicked the adipic section of the succinimidyl ester equivalent synthesis. As a result, the CV-N conjugates formed from each method were not significantly different from each other.

The synthesis of the maleimide-biolinker-toxin was also relatively straightfoward using a solid phase approach. The attachment of the toxin amine *via* the succinimidyl ester was successful with purification by HPLC providing the final product. The two reactions of this toxin construct with the two different forms of thiol containing CV-N were understandably similar and with the halting of the thiolation of CV-N (**15**) at predominantly the mono-thiol stage the products were both primarily a protein with a single attachment of toxin. The difference lies in the 'known' position of the CV-N-Cys toxin while the thiolated-CV-N toxin could be on any one of the lysine side chains, or as is more probable, on several different of the available lysines per CV-N molecule.

Chapter 6

TESTING OF CV-N CONJUGATES

6.1 Introduction

Although the synthesis of the various CV-N constructs was ultimately successful, the determining factor in deciding their therapeutic potential could only be evaluated by extensive biological testing. There were several distinct aspects of the interaction between CV-N (**15**) and gp120 expressing cells that needed to be explored with respect to modifying the CV-N (**15**) skeleton.

The first phase of interaction is the initial binding of the CV-N (**15**) molecule to the gp120 on the outer surface of infected cells. With the attachment of large molecules such as the toxin conjugates to CV-N, there was the potential to block the CV-N binding (**15**) to gp120. This was presumed to be the case for the CV-N-PE38 conjugate previously constructed and tested by Mori *et al.*[108] where the gp120 binding affinity of the conjugate was significantly reduced. The same method of testing using an ELISA system, developed by the LDDRD, was used to determine the relative affinity of the synthesised CV-N constructs for gp120. This assay was critical in determining if the conjugation of compounds to the amines in CV-N resulted in any loss of binding affinity. If a high loading of conjugates onto CV-N resulted in a loss of gp120 binding then this would have to be taken into account in the synthetic strategy with perhaps minimal loading of the CV-N (**15**) protein.

The XTT anti-HIV assay detects drug-induced suppression of viral cytopathic effects by generation of a soluble formazan in surviving cells. For the CV-N-conjugates produced, this analysis tests the binding of the CV-N to the gp120 on the outer surface of the virus. This effectively 'protects' the uninfected cells which therefore continue to metabolise the

XTT to the formazan. Therefore, the results from this assay demonstrate whether or not the CV-N constructs can still productively inhibit the binding of virus to cell even with the addition of the dye molecules.

The first conjugates tested were those from section **4.2** where CV-N was conjugated to two fluorescent dye derivatives under a variety of reaction conditions. Several of these CV-N samples were very heavily loaded with dye molecules and therefore provided a good model for whether conjugated CV-N proteins would still interact effectively with gp120. These dye conjugates were tested using both the ELISA and XTT assays. Further constructs tested were the CV-N-biolinker-coumarin conjugates derived from the succinimidyl ester strategy (section **4.3.6**). These were tested using the XTT assay system to determine if the biolinker derived constructs in any way interfered with the binding of CV-N to gp120. These coumarin conjugates were also 'fed' to productively infected cells to determine if the constructs became internalised into the cell through gp120 binding. Once inside the cell the biolinker should cleave, releasing the free coumarin dye whose fluorescent maxima would change as the amide bond previously linking it to the biolinker would now be an amine. These tests were designed to examine the effectiveness of the biolinker in releasing the coumarin (toxin substitute) internally in infected cells.

The ultimate testing is of course with the CV-N-toxin conjugates themselves to determine if the cell-killing ability of the halichondrin moiety can be selectively applied to productively infected cells. These conjugates have been sent for such biological testing. The results are not yet known.

6.2 Assay Systems

A discussion of the particular assay techniques used for testing is important to clearly understand exactly what properties of CV-N's interactions are being tested. Only in this way can the results be evaluated effectively. The two assays (ELISA and XTT) although both testing the binding of CV-N (or CV-N conjugates) to gp120, do so in very different ways providing different perspectives on the binding process.

6.2.1 ELISA Assay

Immunoassays form the backbone of tests used in the study of infectious diseases such as HIV. The ELISA method provides a highly sensitive and precise method for the estimation of biological parameters with the added advantage of being able to handle large numbers of samples rapidly. Since the first effective enzyme-labelled assay was described in 1971[160] there have been thousands of applications published dealing with the quantification of antigens and antibodies for research.

There are many different categories of ELISA utilising different binding patterns. The particular ELISA for testing CV-N (**15**) binding is based on the 'Sandwich ELISA' with indirect detection of the binding.[161] An outline of the binding process is shown in **Figure 6.1**. The assay is performed in a 96-well plate with initial binding of gp120 to the plate itself. Incubation for two hours is followed by thorough washing of the plate. The resulting sites that are unfilled by gp120 are 'capped' by introduction of Bovine Serum Albumin (BSA) and left incubating overnight. After washing, the CV-N (**15**) or CV-N conjugates are introduced and selectively bind to the gp120. Excess compounds are again washed away and then antibodies introduced that are directed specifically to CV-N (anti-Rabbit-anti-CV-N Ab). These are in turn bound by a second antibody (anti-Rabbit-IgG-Ab-AP) leaving an alkaline phosphatase (AP) enzyme exposed at the terminal. The final step, after again washing, is to add the substrate buffer (10% diethanolamine, 10 mM

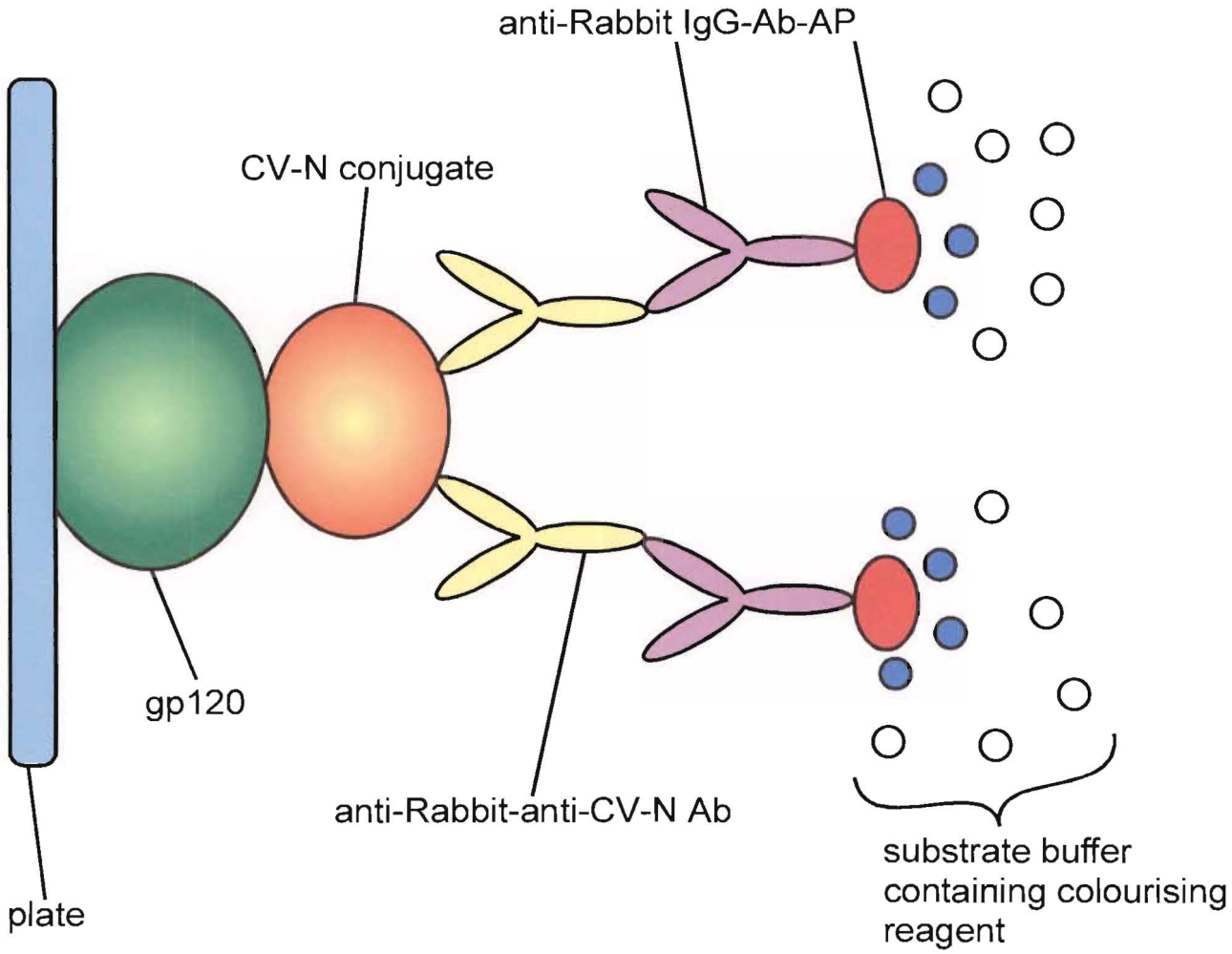


Figure 6.1 ELISA testing of CV-N conjugates

$MgCl_2$, 4 mg/mL *p*-nitrophenyl phosphate (pNPP), pH 9.8). The colour development (due to substrate binding) is halted after 15 min with addition of EDTA. The plate can now be read at 405 nm in a plate reader. Any alterations to the binding of CV-N derivatives to the gp120 will cause fewer antibodies to bind at each step and hence the yellow colour will not be as intense as for the native CV-N (**15**).

A sample plate layout is shown in **Figure 6.2**. The outer ring of wells is loaded with BSA at the same time as the rest of the cells are blocked with BSA after binding of gp120 to the well walls. This gives a baseline reading to ensure the assay is working correctly. Rows B, C and D were used for native CV-N (**15**) as a standard reference for each plate. The concentration of CV-N loaded into each well increases from column 2 to 11 in a

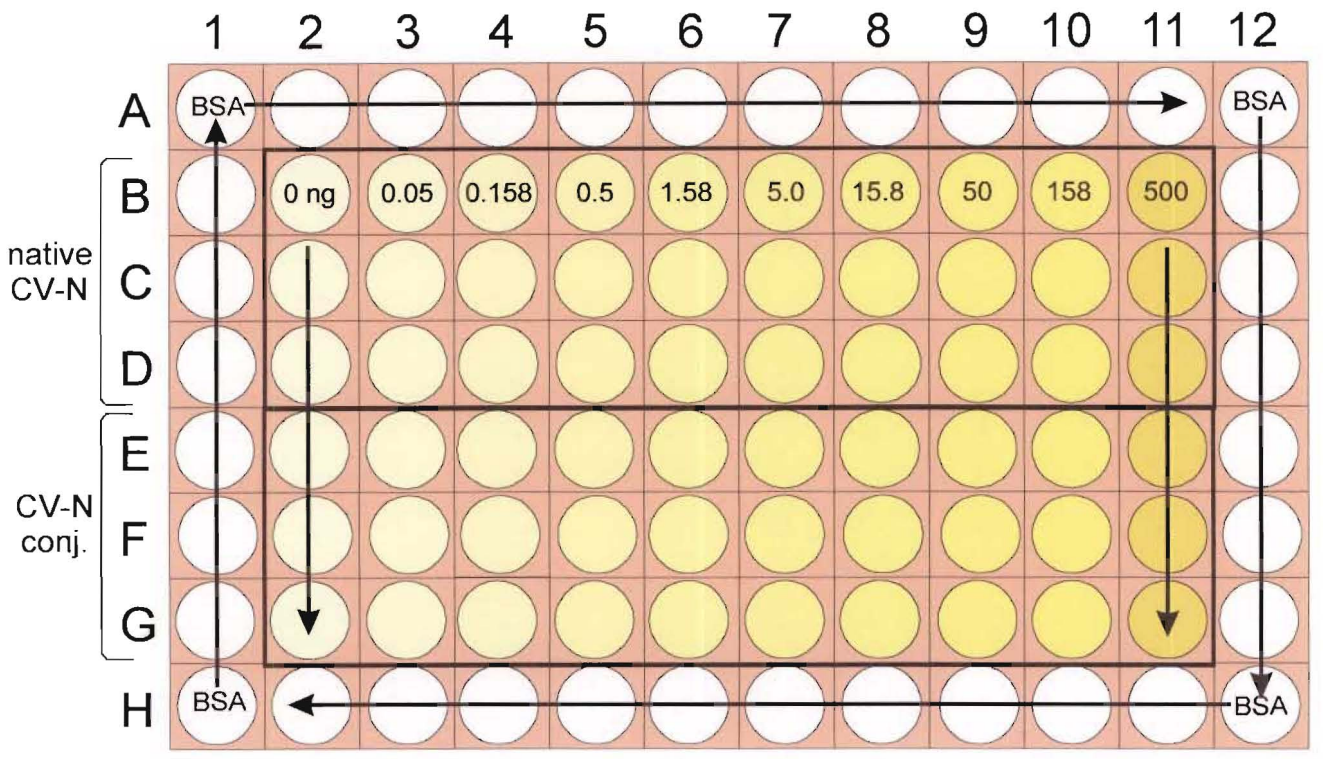


Figure 6.2 Sample ELISA plate

logarithmic fashion by a factor of 3.16 each time. This extensive range of concentrations allows an accurate graphical representation to be presented that gives a precise measure of the binding affinity of the CV-N (**15**) (or CV-N constructs). Each individual concentration of CV-N is assayed in triplicate in order to ensure uniformity of readings and to allow the easy identification of 'erroneous' readings which can occur primarily due to cross-contamination of the wells. A comparison of the data for the native CV-N (**15**) *versus* the CV-N conjugates gives an accurate indication of their relative binding affinities to gp120.

6.2.2 XTT Assay

The particular XTT anti-HIV assay utilised for the analysis of the CV-N-constructs was originally developed in 1989 by the NCI to support and enhance the discovery of anti-HIV agents.[162] Viable cells (those metabolically active) convert the tetrazolium salt XTT into a soluble orange formazan (**Figure 6.3**) therefore giving the culture a high optical density

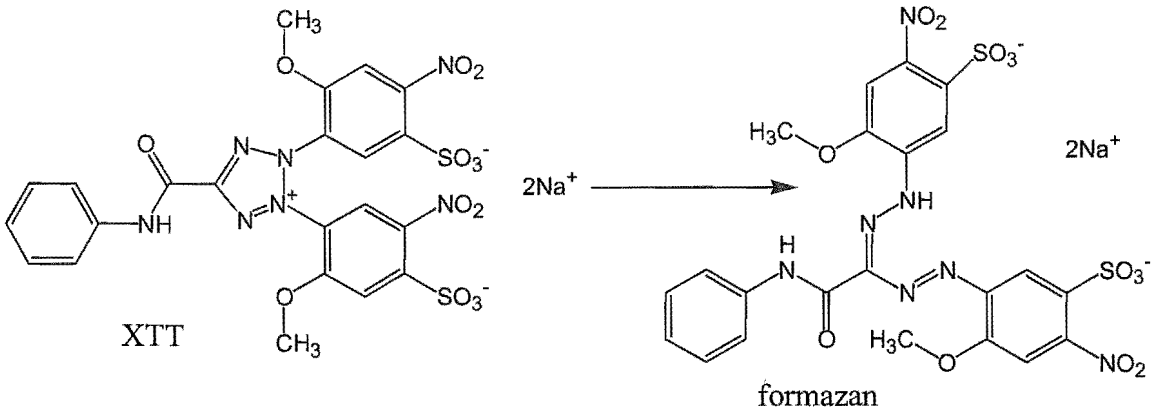


Figure 6.3 Metabolization of XTT to a water soluble formazan salt by viable cells

(OD) at 450 nm. The cells used are lymphocytic CEM-SS cells, the choice of which is based on their sensitivity to the lytic effects of HIV-1 infection and their ability to metabolically reduce XTT. CEM-SS cells that are not protected by drugs such as CV-N (**15**) are killed by the virus and/or do not proliferate, consequently producing less formazan and therefore lower optical densities. This assay is therefore a measure of CV-N's (**15**) ability to 'protect' cells from infection by binding to gp120 and preventing the HIV virus from entering the CEM-SS cells. Results obtained are treated graphically but also produce an EC_{50} value which is the concentration of drug at which 50% 'protection' of the CEM-SS cells is obtained. The XTT assay also measures any direct cytotoxic effect the drugs tested may have to the CEM-SS cells in the absence of virus.

The protocol is to initially dilute each drug in complete medium into a single column of a 96 well plate (dilution plate). An automated workstation then performs eight serial dilutions of each drug and transfers 100 µL of each dilution to the test plate. Uninfected CEM-SS cells are then plated at a density of $1x10^4$ cells in 50 µL of complete medium. Diluted virus is then added to the appropriate wells in 50 µL to yield a total volume of 200 µL in each well. After addition of the cell, virus, and drug controls, the plate is incubated at 37^0C in an atmosphere containing 5% CO_2 for 6 days. Then 50 µL of the XTT solution is added to each well of the test plate and incubated for 4 hours at 37^0C. After incubation,

the plates are covered with adhesive plate sealers, shaken, and their optical densities determined at a test wavelength of 450 nm.

A sample plate layout for this assay is shown in **Figure 6.4**. Two drugs are tested per plate (shown as blue and green) with rows B to E containing the active virus to test the infectivity. The cytotoxicity of the drugs to the CEM-SS cells is tested in rows F and G with standards of the drugs in the absence of cells or virus, in rows A and H. The standards of both CEM-SS cells and the virus are present on each plate to ensure the assay is performing correctly with viable cell growth and viral activity. Much like the ELISA assay, the use of 96 well plates allows extensive automation and the screening of large quantities of samples with high reproducibility.

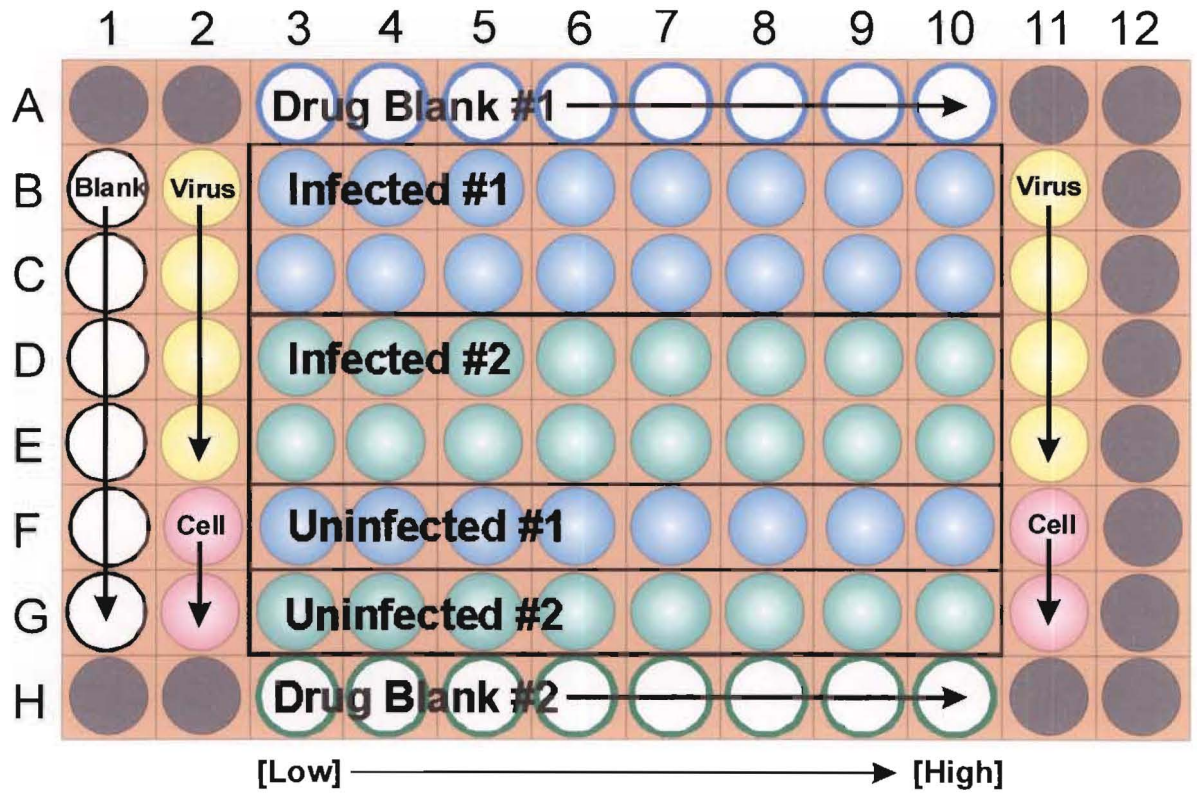


Figure 6.4 Sample XTT plate

Figure 6.5a) shows a light microscope image (50x) of how the CEM-SS cells appear in the absence of virus. The cells are small and well defined with the vivid yellow colour reflective of the stain used to visualise them. The second image (**Figure 6.5b)**), shows the same CEM-SS cells after HIV virus has been introduced and the cells incubated for 7 days (as per the XTT assay protocol). The overwhelming formation of virus-induced multi-nucleated giant cells (syncytia) and lack of small, well-defined cells is immediately evident. The HIV-induced cytopathic effects eventually result in the destruction of syncytia leading to suppression of both virion synthesis and formazan formation. A closer examination of these cell types is shown in **Figure 6.5c)** and **d)**. The magnification has increased from 50x to 200x and the individual CEM-SS cells in **c)** can be seen more clearly.

The protective effect CV-N (**15**) imposes on these CEM-SS cells is shown in **Figure 6.6a)-b)**. Both pictures have had the same treatment with equal amounts of virus added to them and subsequent incubation for 7 days. There is no evident scyntium formation in the 50 nM CV-N (**15**) treated cells in **a)** while the beginnings of multinucleated cells (as indicated by arrows), can be seen in **b)**, at 0.5 nM concentration of CV-N (**15**). Without the presence of CV-N (**15**) these pictures would appear as the cells in **Figure 6.5b)**. The question regarding the CV-N-conjugates under investigation was whether or not the alterations made to the CV-N protein affected its binding, and hence protective qualities, in this XTT assay.

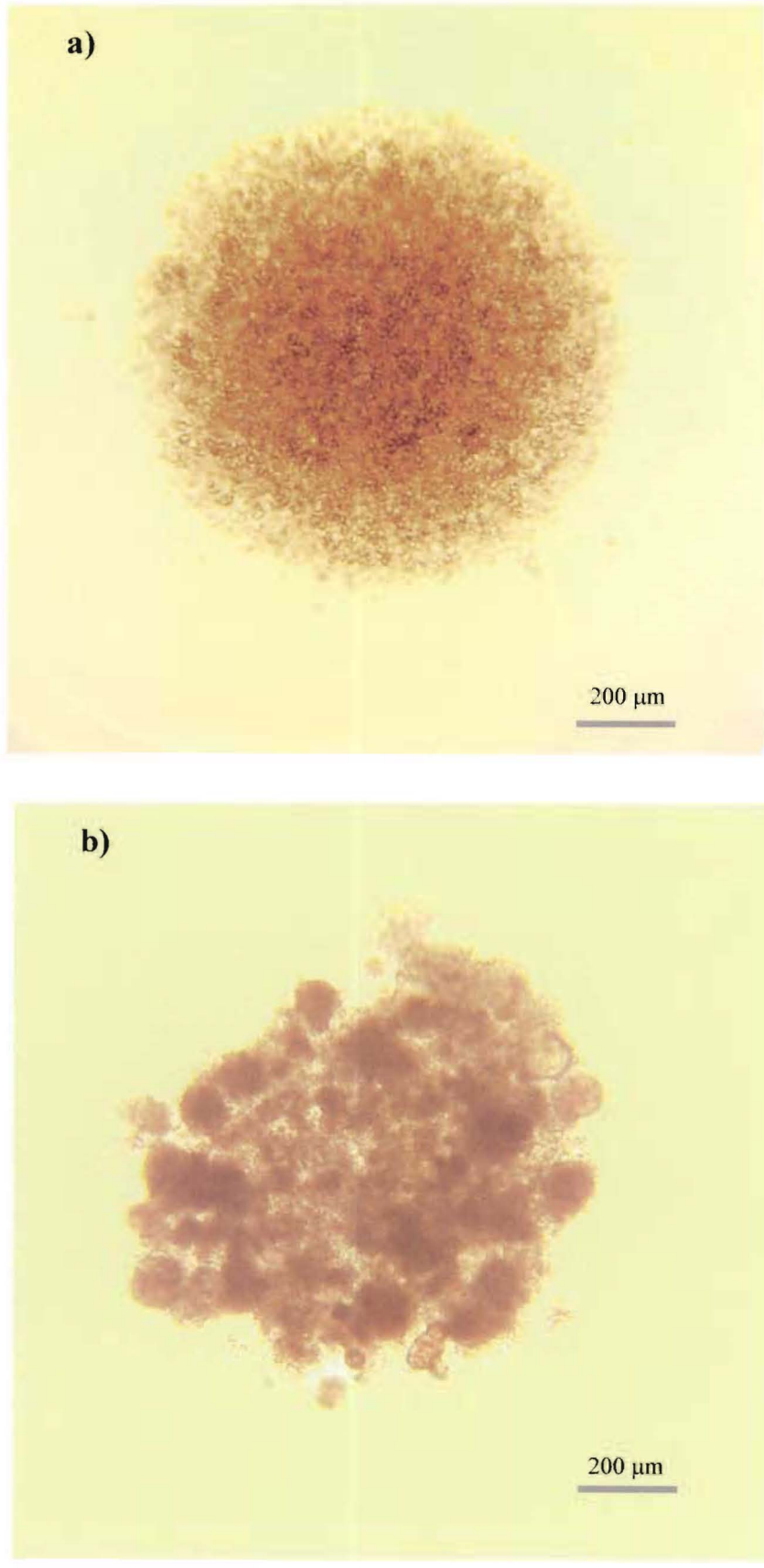


Figure 6.5a)-b) Light microscope image (50x) of **a)** uninfected and **b)** infected CEM-SS cells

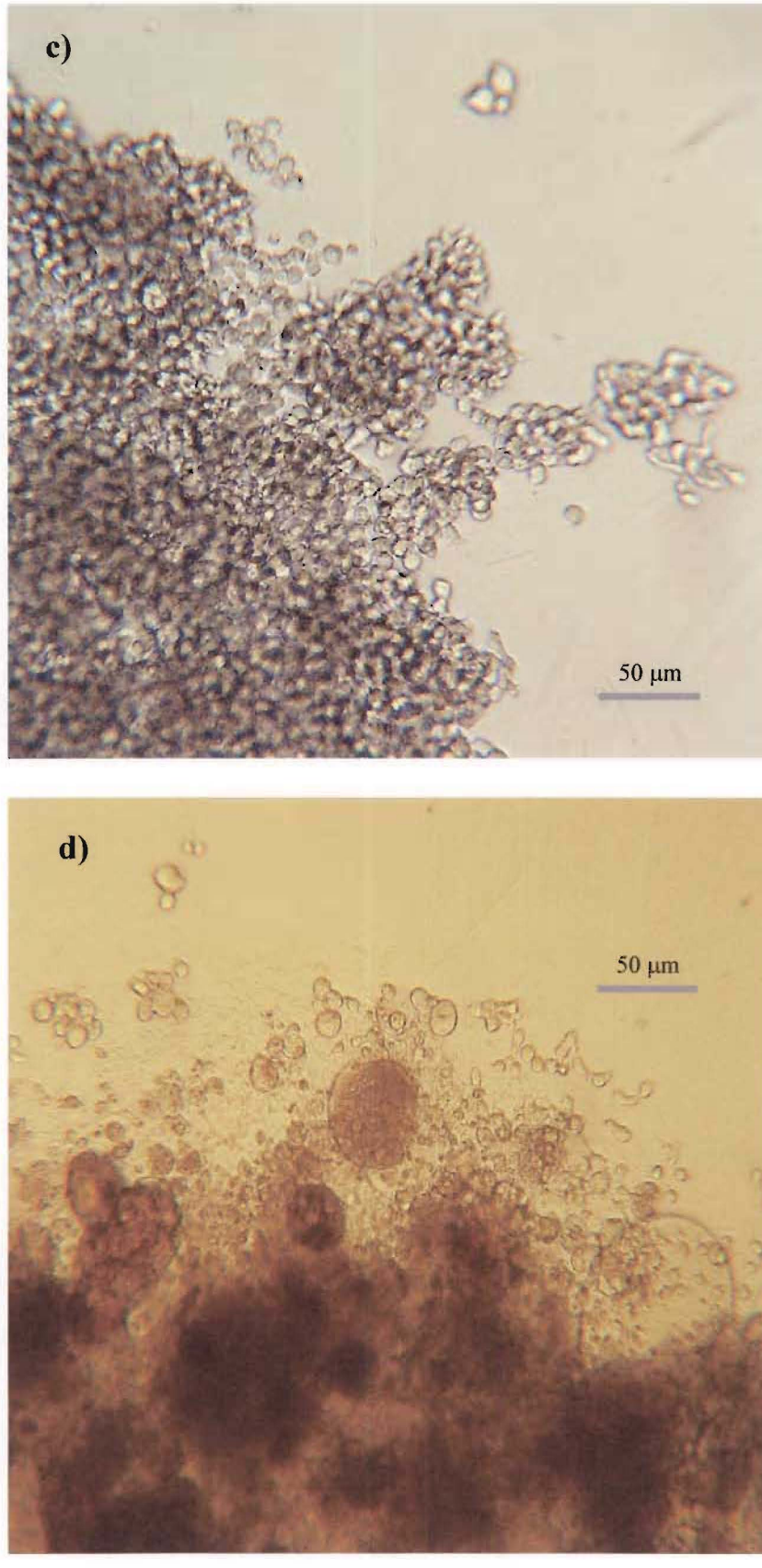


Figure 6.5c)-d) Light microscope image (200x) of **c)** uninfected and **d)** infected CEM-SS cells

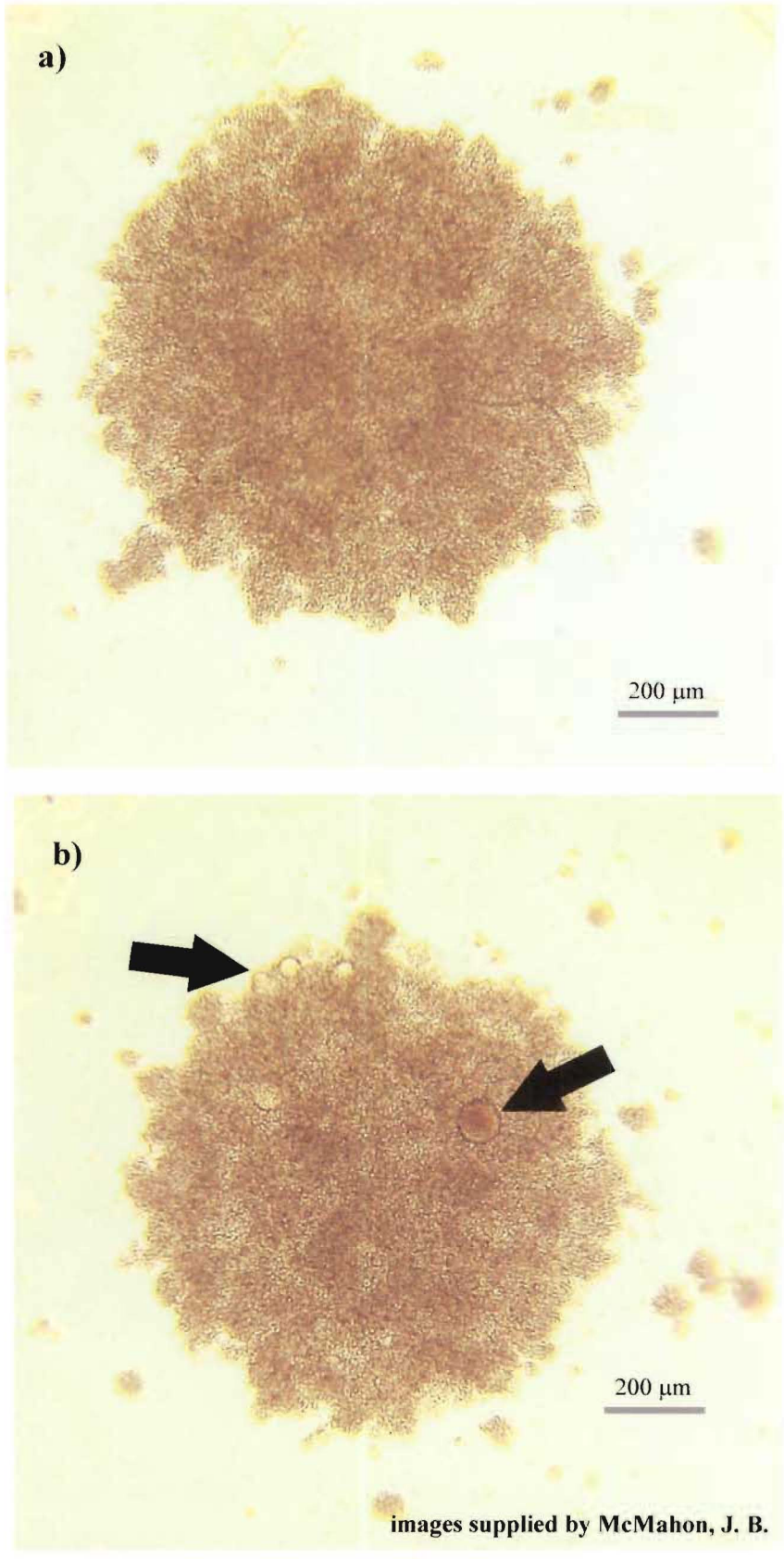


Figure 6.6a)-b) Light microscope image (50x) of **a)** CV-N (50 nM) and **b)** CV-N (0.5 nM) protected CEM-SS cells

6.3 Testing of CV-N-Dye Conjugates

6.3.1 Initial ELISA Testing

Before the CV-N-fluorescent dye derivatives could be tested effectively by the ELISA assay system the effect of the dye moiety (if any) on the assay system itself needed to be investigated. This was to prevent false positive or false negative results from influencing the analysis of the data. There were two issues in particular, the first of which was whether the dye interfered with the binding of the antibodies, and secondly, whether the dye contributed to the substrate reading at 405 nm. The first of these problems would give a false negative result as the CV-N-conjugates may bind to gp120 adequately but if the antibodies did not bind, then there would be no response at 405 nm as the substrate would not have bound either. The second problem would be more subtle in that the result would still be positive although the intensity would be increased due to the 'addition' of the dye chromophore to the substrate response. This may lead to attributing a stronger binding affinity to the conjugates than would be strictly true.

The first issue was tested by performing the ELISA assay without the presence of gp120. The first step was effectively coating the wells with the CV-N (**15**) (or CV-N-conjugate) in an increasing concentration. Without the gp120, the CV-N conjugates would bind to the well regardless of their 'loading' of dye. The remaining space in the well would be filled by BSA as for the standard ELISA. The antibodies would also be introduced as per the normal procedure. It was not necessary to test all the dye conjugates so only the most highly loaded samples for each dye from the pH 8.3 series were used (**36** and **42**). [At this stage it was not known that the BODIPY conjugate chosen for this test (**42**) was one of the samples that had not performed as expected, in that very little protein material was present.] It was found that the CV-N-conjugates (particularly **36**) did not interfere with the antibody binding as the results for the CV-N conjugates were comparable to the results obtained for the native CV-N (**15**) in the same procedure. Therefore, the presence of the

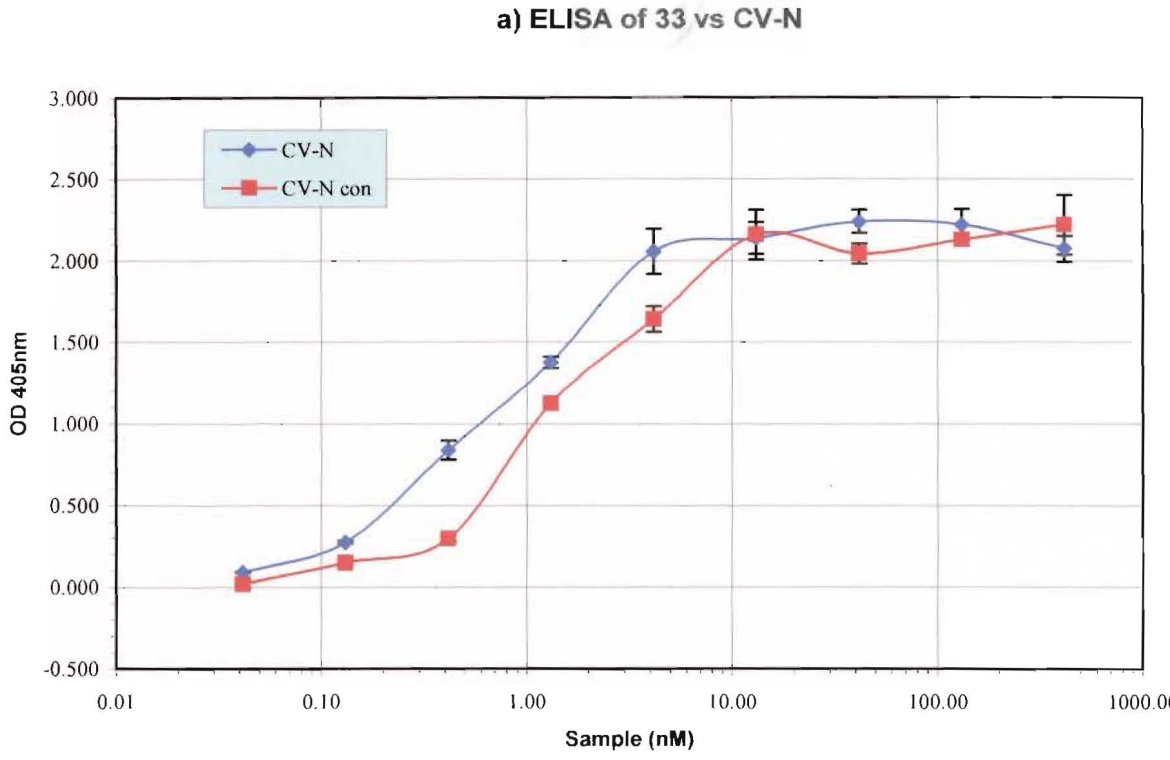
dye moieties did not prevent the antibodies from performing their task in binding to CV-N and ultimately producing a substrate response.

The second issue, contributions to the absorption at 405 nm from the fluorescent tag, was also tested in the same experiment. This was done by reading the plates in the standard ELISA plate reader at 405 nm before addition of the substrate buffer. Without the buffer there should only be baseline response from any of the wells at 405 nm unless the dye was itself contributing to the absorption. There was no discernible detection of absorption at this wavelength for either dye tested even at the highest concentrations. This meant that the dye moieties would not interfere in any way with the response from the substrate in the ELISA experiment.

As a check to determine if in fact there were samples of each dye conjugate present in the wells in increasing concentrations, the plates were exposed to a wavelength close to the absorption maxima of the respective dyes. The plate reader then recorded any response at the emission wavelength for the dyes. As the concentration of the conjugates increased across the plate, so did the fluorescent response as expected. This could also theoretically be used to determine the binding affinity of the fluorescent CV-N conjugates to gp120 if appropriate controls and standards were in place.

6.3.2 ELISA Testing of 5-SFX and BODIPY CV-N-Conjugates

Each of the mixtures of constructs synthesized during the reactions at pH 8.3 in section **4.2** were tested using ELISA. [The samples from reactions at pH 7.1 were not tested as many contained a high proportion of native CV-N (**15**) and an ELISA assay would reveal little data of value.] Each sample tested was loaded onto a 96 well plate with native CV-N (**15**) also present as a control. The data from the response of the detection substrate at 405 nm has been converted into graphical form in **Figure 6.7** for the CV-N-5-SFX conjugates and **Figure 6.8** for the CV-N-BODIPY conjugates. With the exception of the 25 equiv.



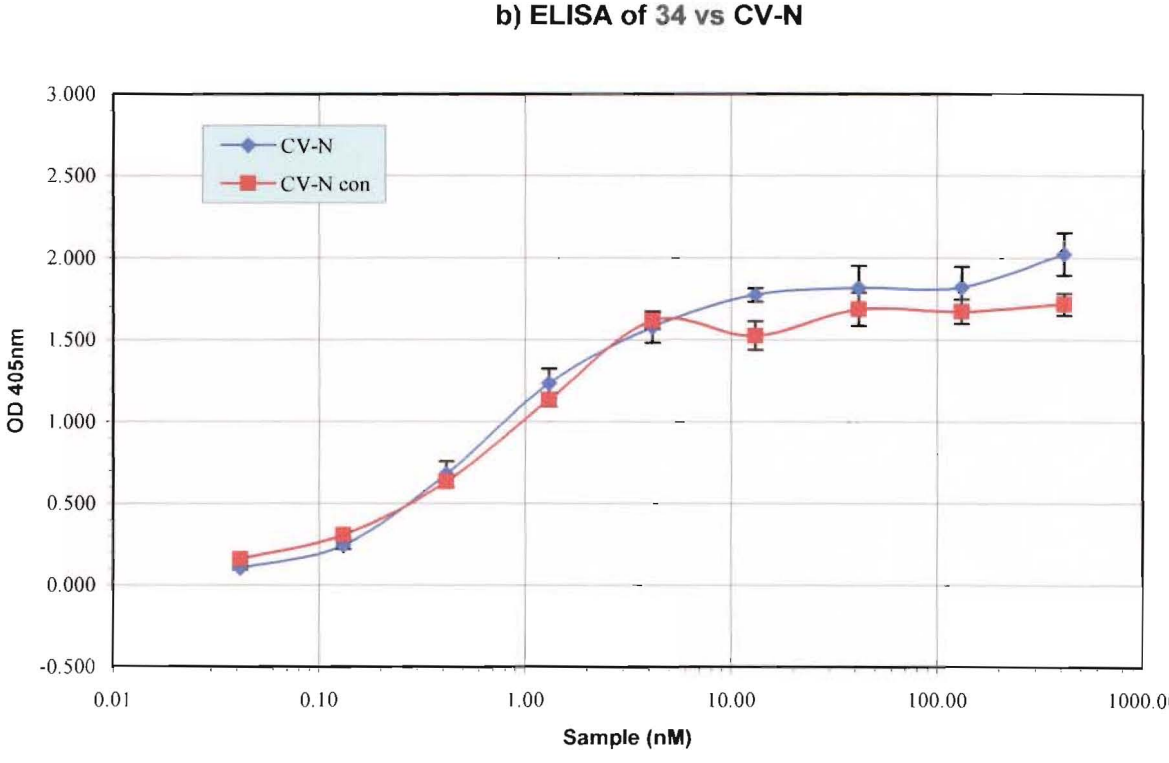
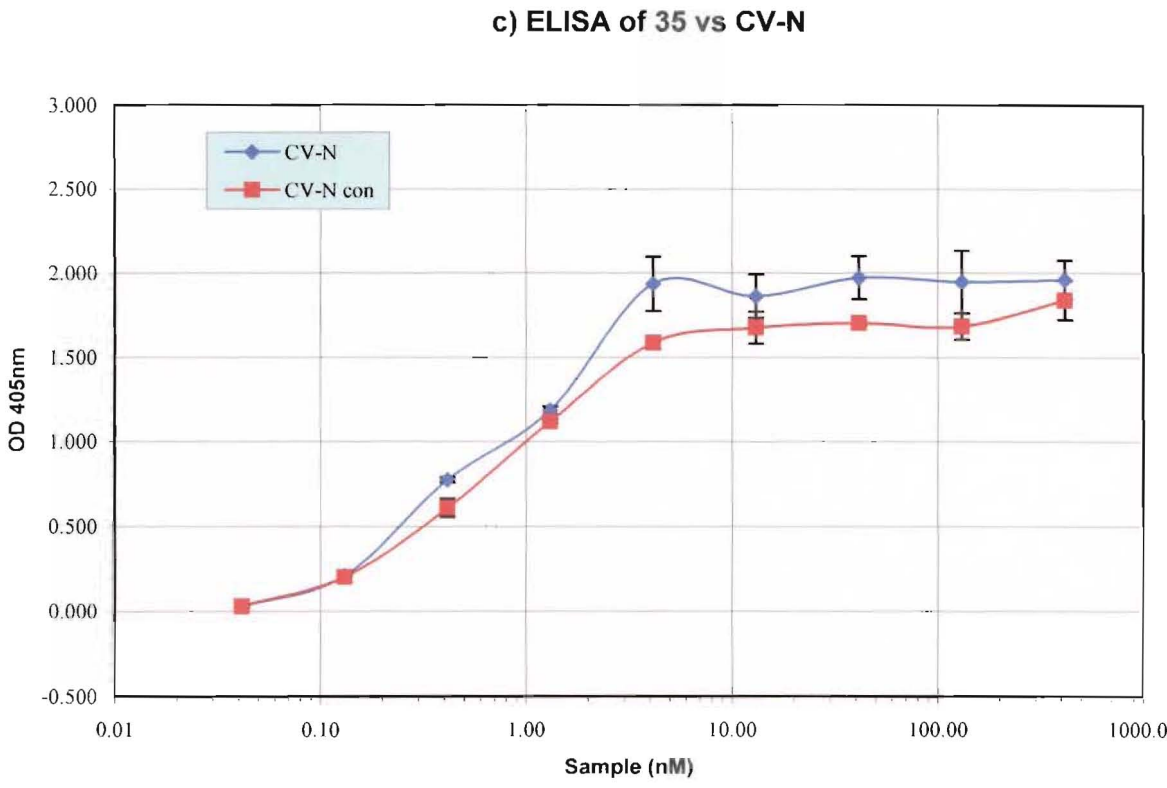


Figure 6.7a)-b) ELISA of CV-N (**15**) and CV-N-5-SFX conjugates



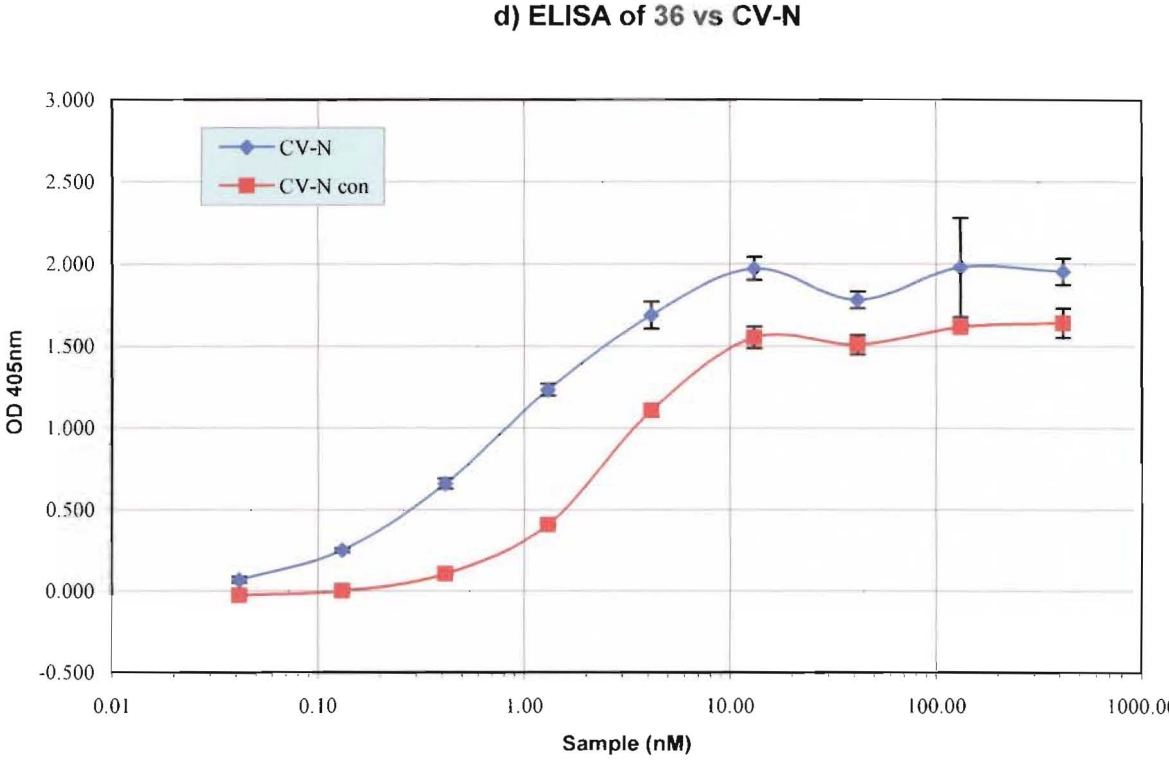
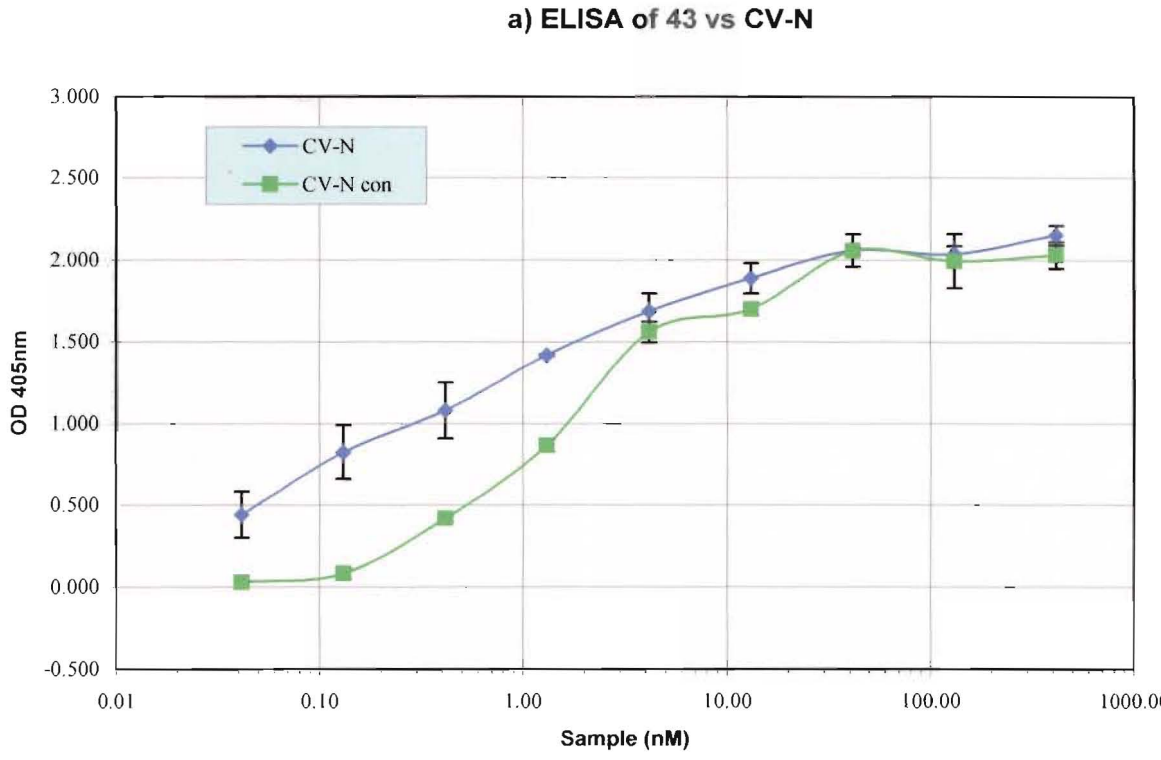


Figure 6.7c)-d) ELISA of CV-N (**15**) and CV-N-5-SFX conjugates



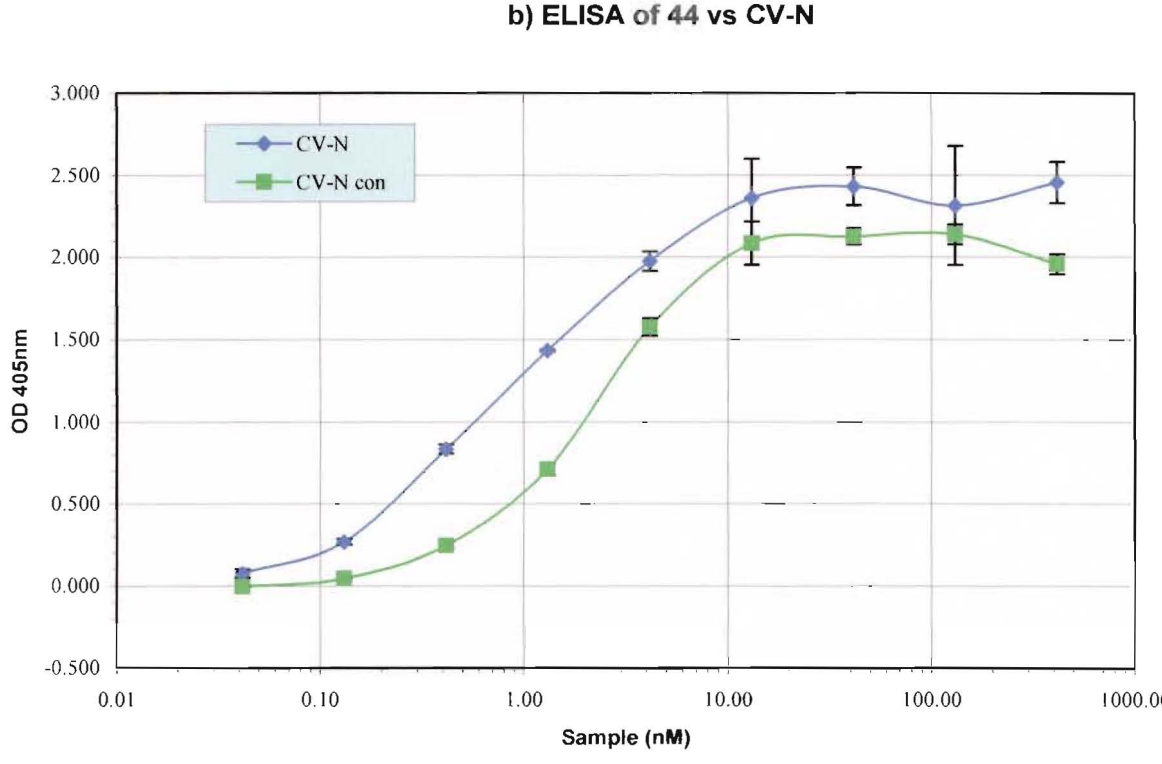
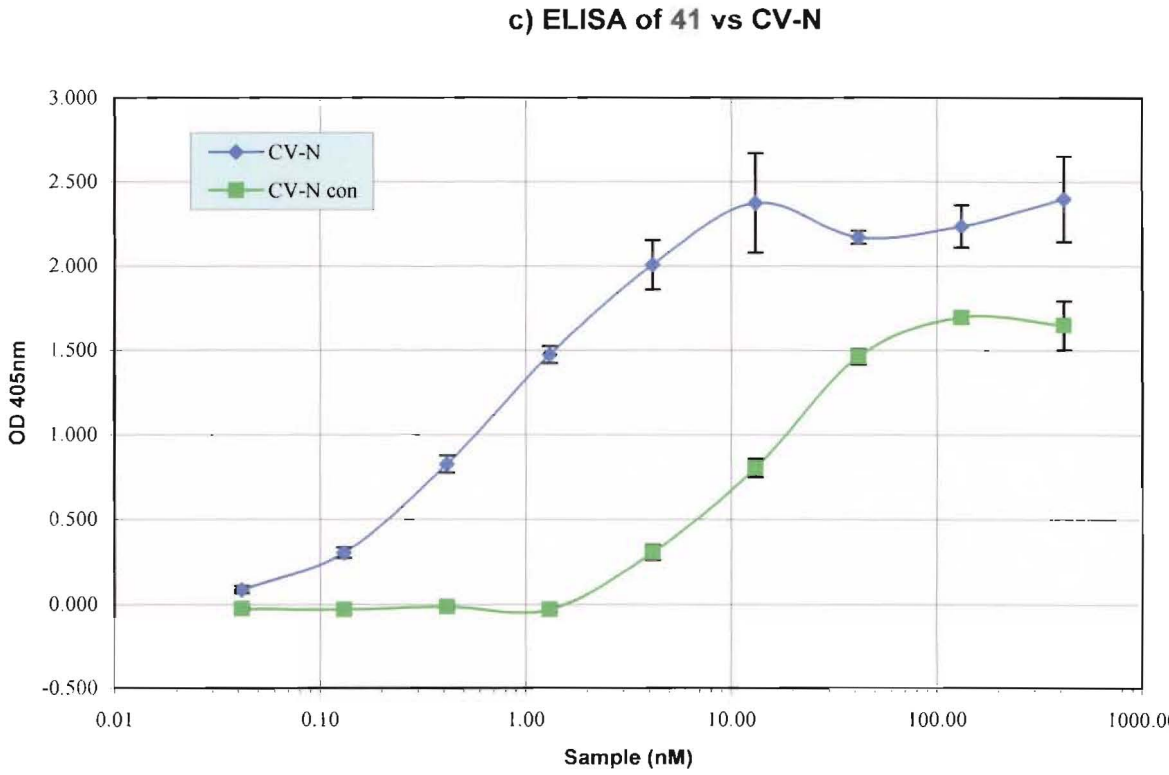


Figure 6.8a)-b) ELISA of CV-N (**15**) and CV-N-BODIPY conjugates



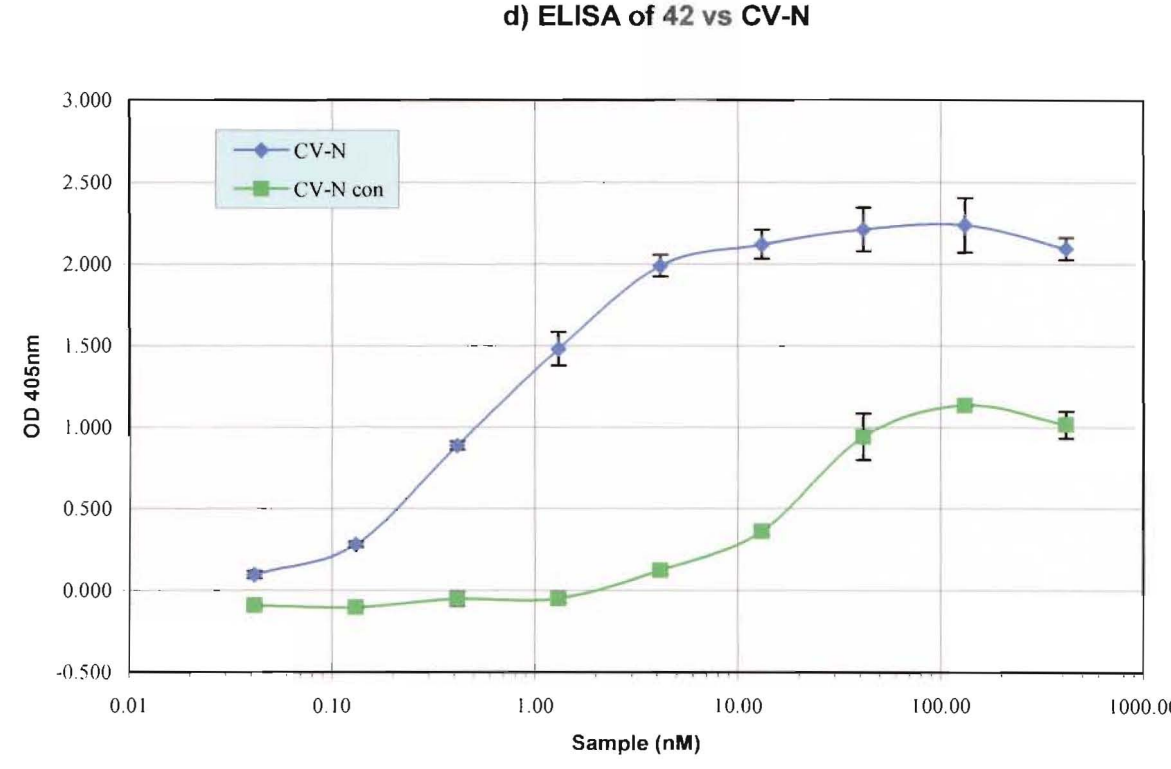


Figure 6.8c)-d) ELISA of CV-N (**15**) and CV-N-BODIPY conjugates

BODIPY conjugates (**Figure 6.8c)** and **d))** the remainder of the samples can clearly be seen as comparable to native CV-N (**15**). As the sample concentration increased so did the substrate response as the amount of bound CV-N-conjugate to gp120 had also increased. The binding affinities of **41** and **42** are markedly reduced due to the fact that for some undetermined reason the CV-N (**15**) (or conjugates) were not present in the fraction collected from the G25 purification. They do show some binding affinity at higher concentrations which may indicate the presence of smaller concentrations of CV-N (**15**) (or conjugates) that were undetected by LCMS analysis.

Some samples still contained a reasonable percentage of native CV-N (**15**) and therefore one would expect that they would show binding affinities comparable to that of pure native CV-N (**15**). However, the samples with very little or no native CV-N (**15**) remaining and a high loading of fluorescent tags, also show binding comparable to the native protein. The precise position of the binding site(s) on gp120 and the corresponding region(s) on CV-N (**15**) have not yet been accurately determined. Preliminary findings tend to suggest that the binding sites are distant from the lysine residues but there remains a large degree of uncertainty. The results of O'Keefe *et al.*[78] showed that the required ratio of CV-N (**15**) to gp120 was 5:1 so there may be a variety of sites on CV-N (**15**) that bind to pockets on gp120. With 'blocking' of some sites by conjugation of the lysine side chains, the CV-N conjugate may bind at another, more accessible, site. Whatever the position of binding, the overall conclusion from these particular studies was that attaching relatively small molecules (MW ~450 Da) to the lysine side chains (or possibly the *N*-terminal) did not interfere with the binding affinity of CV-N for gp120.

6.3.3 XTT Assay Testing of CV-N-Dye Conjugates

The concentrations of each of the CV-N-dye conjugates in solution differed. This was a consequence of differing volumes collected from the G25 column in each case during purification (section **4.2**). The samples were each submitted for the XTT assay as 200 μL aliquots from each of the eight conjugates tested (four for each dye at pH 8.2). Therefore,

the concentrations of each sample in the wells of the test plates differed from each other making a comparison of the data misleading. However, the EC_{50} values obtained were scaled according to the concentration of each sample and it is these values that can be directly compared with each other and with native CV-N (**15**) itself. Note that **Figure 6.9** does show the raw data obtained from the most heavily loaded CV-N conjugate (**36**).

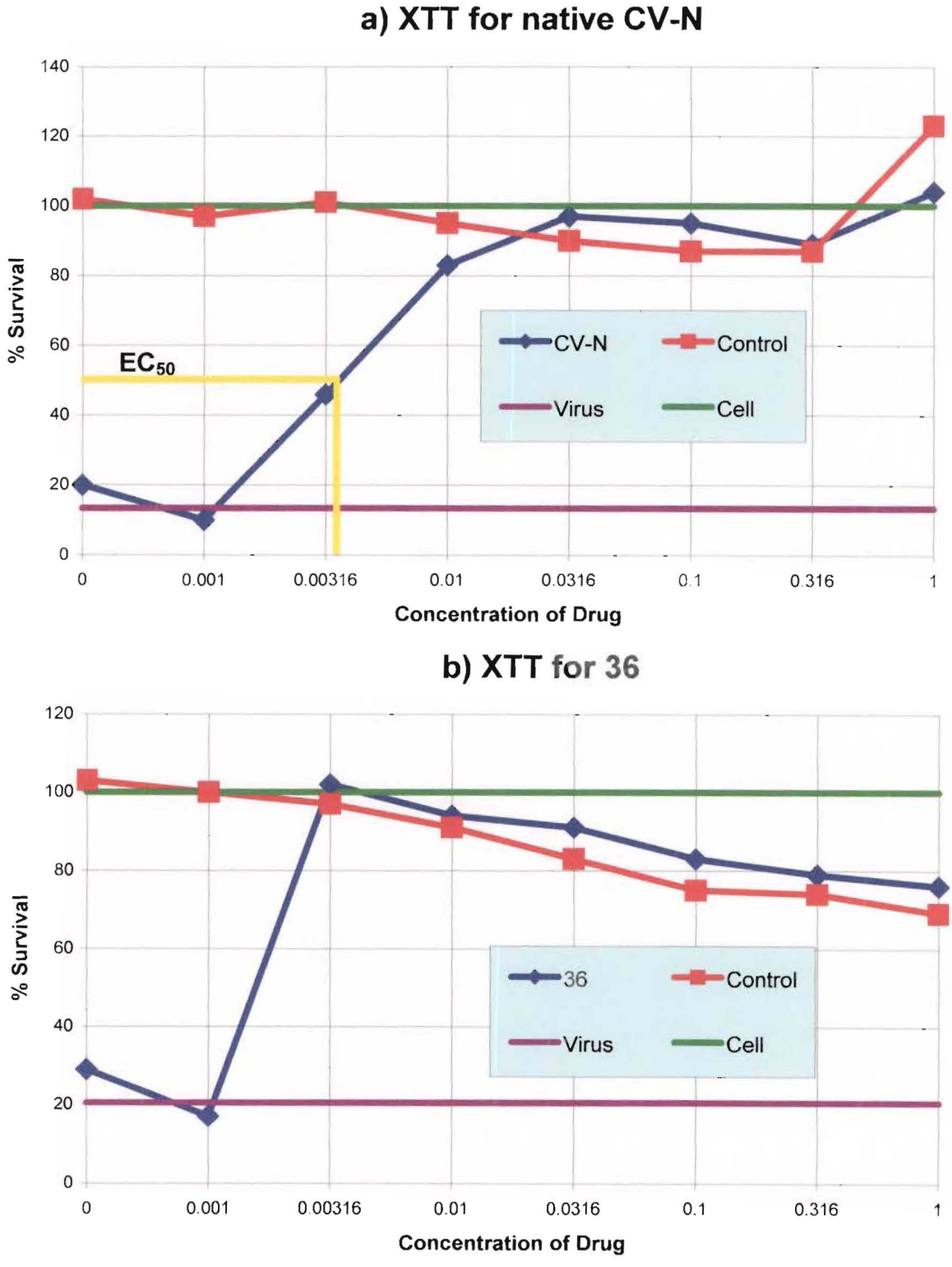


Figure 6.9 XTT assay results for **a) 36** and **b)** CV-N (**15**)

This data is shown in comparison to that of native CV-N (**15**) itself in order to show how the analysis took place. It should be carefully noted that sample **36** was almost exactly twice as concentrated as the native CV-N (**15**) sample. The data representing the 100% cell survival (green line) was obtained from an average of the standard cell blanks in the 96 well plate. In contrast, the line representing the percent survival of cells completely infected by virus (purple line) without protection from any drugs was approximately 20% for sample **36** and 13% for native CV-N (**15**). The uninfected cells gave an approximately constant measure of cell viability as the concentration of the CV-N (or adduct) increased. The data for the infected cells started with a low percentage survival at a low concentration of CV-N (**15**) as there was no cell 'protection'. As the concentration increased, so did the protection afforded by CV-N (**15**) with a concomitant increase in percentage survival of the CEM-SS cells. Due to the higher concentration of **36** present (compared to native CV-N (**15**)) the line representing the infected cells for **36** climbed at an earlier concentration than that for the native CV-N (**15**).

The EC_{50} value was obtained by correlating the 50% percent survival mark with the concentration as shown for CV-N in **Figure 6.9a)**. The EC_{50} values for the remaining samples were obtained in the same way. The corrected EC_{50}'s (i.e. taking into account the different concentrations of the CV-N-dye conjugates) are shown in **Table 6.1**.

Table 6.1 EC_{50} results from XTT assay of CV-N-dye conjugates

Compounds	EC_{50} μg/mL
5-SFX, 2eq, 15min (**33**)	0.0026
5-SFX, 6eq, 15min (**34**)	0.0085
5-SFX, 25eq, 15min (**35**)	0.0054
5-SFX, 25eq, 60min (**36**)	0.0034
BODIPY FLX, 2eq, 15min (**43**)	0.0064
BODIPY FLX, 6eq, 15min (**44**)	0.0110
BODIPY FLX, 25eq, 15min (**41**)	0.0897
BODIPY FLX, 25eq, 60min (**42**)	1.2133
native CV-N	0.0040

The values for **41** and **42** were reflective of the lack of any CV-N type compound present, as the interpolated EC_{50} values were far above the value for native CV-N (**15**). The balance of the EC_{50} values, representing those compounds with CV-N partially and fully loaded with dye molecules, were within a factor of three of the EC_{50} value for native CV-N (**15**). With respect to the parameters of the XTT assay these results were very close and it can justifiably be claimed that there was no difference between the protection afforded by the CV-N conjugates or native CV-N (**15**) itself. This was further evidence that confirmed the ELISA assay results suggesting alterations to the amines of CV-N (**15**) did not inhibit the anti-HIV activity of the CV-N itself.

6.4 Testing of CV-N-Coumarin Conjugates

The CV-N-coumarin conjugates synthesised in section **4.3.6** (**58**, **59**) took the testing of the overall strategy a step further. They each contained the cleavable biolinker and so allowed a direct test of this facet of the strategy. Both compounds produced in section **4.3.6** were initially tested in the XTT assay to confirm that they bound adequately to gp120 and protected CEM-SS cells from infection.

6.4.1 XTT Assay Testing of CV-N-Coumarin Conjugates

Unlike the CV-N-dye conjugates these CV-N-biolinker-coumarin adducts were both submitted at the same concentration as the native CV-N (**15**) allowing a direct comparison of the raw data (**Figure 6.10a)-c)**). The trend shown for the infected cells (blue line) in the native CV-N (**15**) result (**a)**) was closely mimicked by that of **58** (**b)**) and **59** (**c)**). This

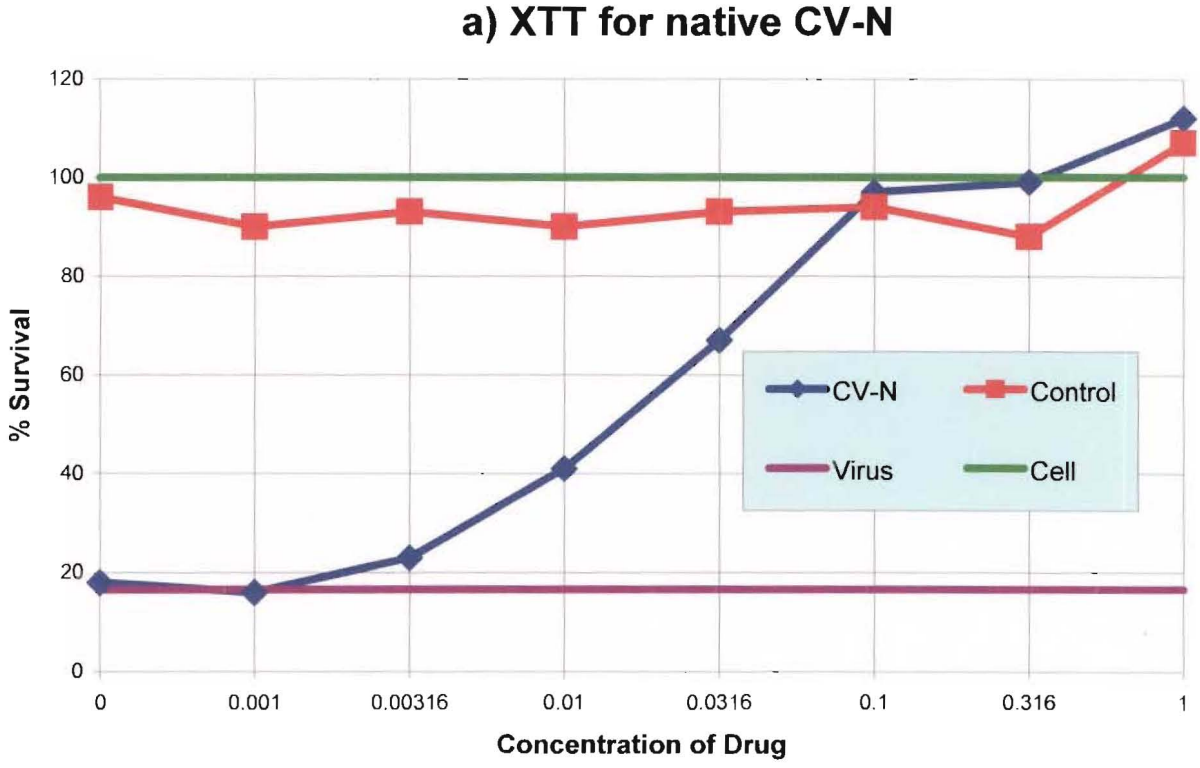
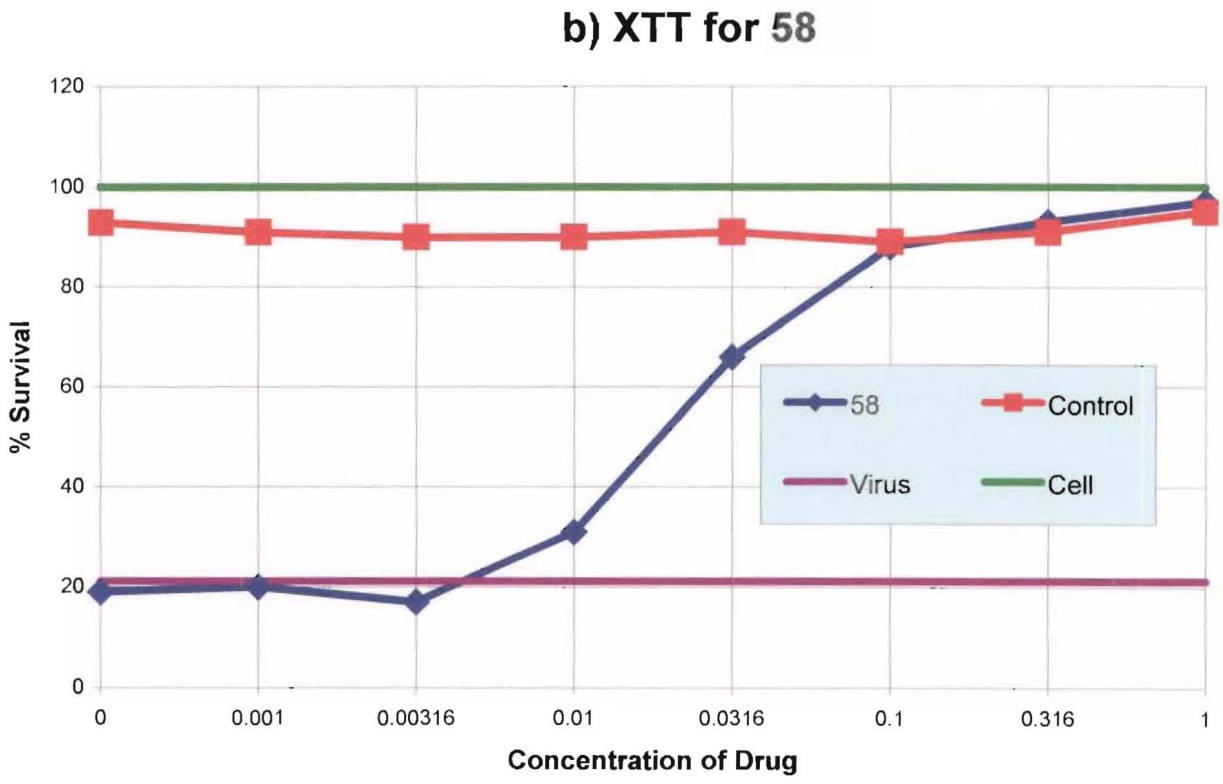


Figure 6.10a) XTT assay results for native CV-N (**15**)



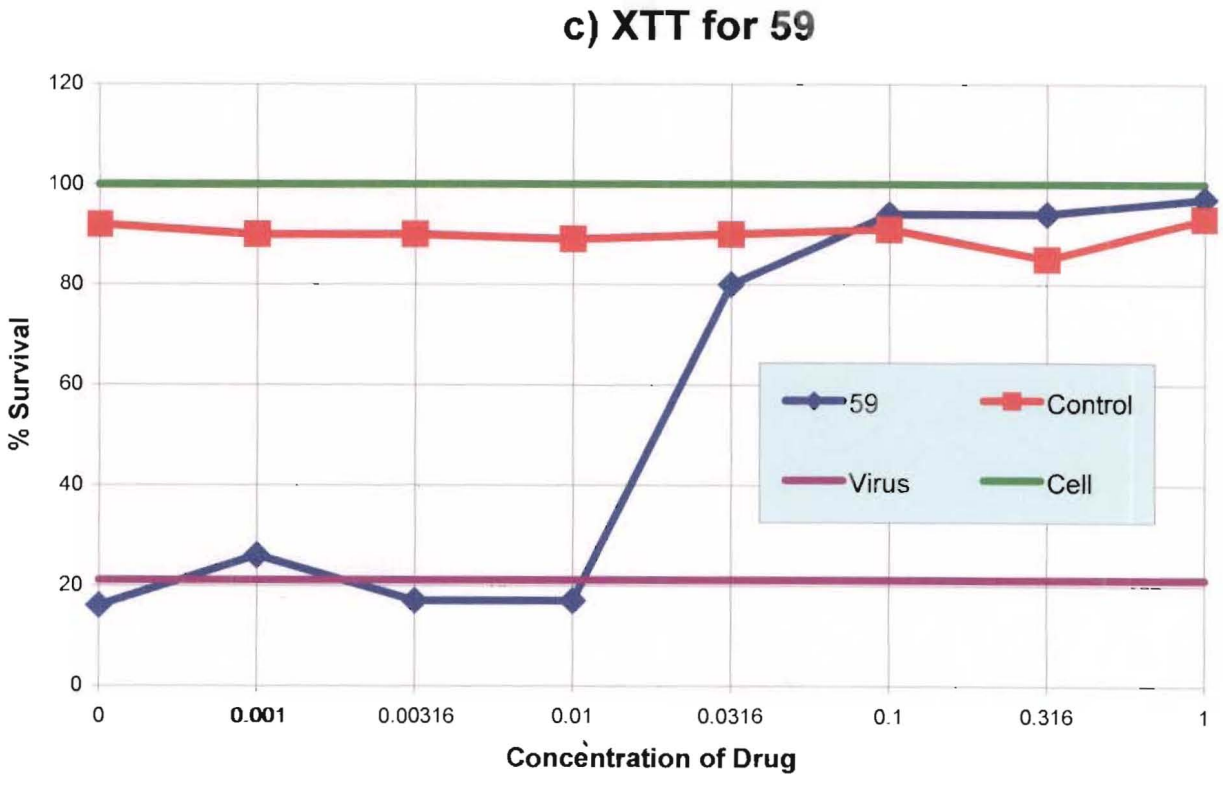


Figure 6.10b)-c) XTT assay results for **b) 58** and **c) 59**

indicated that they both bound to gp120 and both protected the CEM-SS cells just as the native CV-N (**15**) did. The EC_{50} results (**Table 6.2**) reflected how very close the two compounds tested were to CV-N (**15**) itself. The EC_{50} value obtained for CV-N (**15**)

Table 6.2 EC_{50} results from XTT assay of CV-N-coumarin conjugates

Compound	EC_{50} μg/mL
CV-N-Biolinker-Coumarin 15 min reaction (**58**)	0.0210
CV-N-Biolinker-Coumarin 60 min reaction (**59**)	0.0200
native CV-N	0.0160

from this XTT assay was higher than that for the previous CV-N sample (**Figure 6.9** and **Table 6.1**). Inevitably, there is variability in the assay from day to day as fresh CEM-SS cells may not have grown to exactly the level of a previous assay, or a multitude of other minor changes may have slightly affected the result. The controls in place ensured that these changes were within tolerable limits. An increase of 4 times the EC_{50} is not statistically significant and the results can be considered identical. This result stressed the importance of incorporating a control such as CV-N (**15**) for each assay submission in order to have a reference point for analysing the performance value of the samples. This means that only the data in each submission can be accurately compared although obvious trends are evident through multiple XTT assay submissions.

The conclusion from these successful results was similar to that for the CV-N-dye conjugates. That is, the addition of the biolinker-coumarin construct to the amines of CV-N appeared to have had no effect on the binding affinity and subsequent anti-HIV activity of CV-N. In particular, these results went a step further in that they also showed that the adipic-biolinker section of the construct in no way interfered with the binding of CV-N to gp120. This was important for the overall strategy as the only difference between these conjugates and the final toxin derivatives was the substitution of the coumarin moiety by the haliamine toxin. There is an obvious increase in size between the coumarin and the haliamine toxin although the separation from the surface of the CV-N is considerable due to the adipic-biolinker portion. It was decided not to test these CV-N-coumarin conjugates

using the ELISA assay as the results would undoubtedly have shown acceptable levels of gp120 binding as had been noted with the XTT assay.

6.4.2 Further Testing of CV-N-Coumarin Conjugates

The prime reason for synthesising the coumarin derivatives was to establish that the biolinker portion of the constructs would in fact cleave under *in vivo* conditions. The cleavage of the biolinker should release free 7-amino-4-methylcoumarin (**50**) from the conjugate. The CV-N-coumarin conjugate was added to gp120-producing, productively infected CEM-SS cells. In theory the CV-N portion should bind to the expressed gp120 (just as in the XTT assay) with the whole conjugate subsequently taken into the cell as the receptors are recycled.[135] This would in turn lead to enzymatic cleavage of the biolinker at ~pH 5.5 in the lysosomes of the cells, releasing the coumarin amine into the lysosome where it can now diffuse into the cytoplasm. The coumarin when conjugated to the biolinker in an amide bond has an absorption maxima around 324 nm and an emission maxima at 390 nm in aqueous solution. In contrast to this, the amine absorbs at 342 nm and emits at 441 nm. Therefore, as the coumarin dye is cleaved from the conjugate the fluorescence at 441 nm would increase to reflect this cleavage.

However, when this experiment was performed, no fluorescence was observed at the emission wavelengths for the conjugated or for the free coumarin. Unhydrolysed coumarin (conjugated) molecules are known to only fluoresce weakly[149] and it was possible that the presence of the CV-N and biolinker had quenched the fluorescence of the dye. Also, aromatic amines such as 7-amino-4-methylcoumarin (**50**) are partially protonated at low pH (~5) and this may also contribute to the absence of fluorescence with the pH of the lysosomes being around pH 5.5. In retrospect, a fluorescent dye with a larger extinction coefficient should have been used to provide for more sensitive detection to unequivocally establish that the biolinker can be cleaved by gp120-producing, productively infected CEM-SS cells.

6.5 Testing of CV-N-Toxin Conjugates

Three different CV-N-toxin conjugates were synthesised. The first was derived from the succinimidyl ester strategy by coupling to the lysine side chain amines in CV-N (**15**). The second was made from reaction of the maleimide-derived toxin construct coupled to CV-N-Cys. The last was reaction of the same maleimide-toxin construct with thiolated-CV-N. The overall structure of each was very similar with primarily a single attachment of the halichondrin toxin per CV-N molecule. It was unfortunate that the CV-N-coumarin conjugates did not reveal evidence about *in vivo* cleavage of the biolinker due to lack of a fluorescent response. The cleavage of the linker is vital as it is necessary to release the norhomohalichondrin B amine (**28**) in order to kill the infected cells.

The testing these toxin conjugates will undergo is very similar to that performed on the CV-N-PE38 conjugate produced by Mori *et al.*[108] The conjugates will be fed to separate batches of uninfected H9 cells and H9/HIV-1_{IIIb} cells chronically infected with HIV-1 in a modified XTT type assay. The infected cells (gp120 expressing) should be selectively targeted and killed by the CV-N-toxin conjugate while the uninfected cells would remain viable. Therefore, there should be a clear distinction between the two types of cell in terms of the XTT assay. The values derived from the data will be IC_{50}'s which are the effective concentration at which 50% of cells are killed (or disabled). The IC_{50} for the infected cells should be much lower than that for the cells that do not express gp120. This is what was found for the CV-N-PE38 conjugate.[108]

In a similar assay, H9/HIV-1_{IIIb} cells will be treated with the CV-N-toxin conjugate in the presence or absence of free native CV-N (**15**). The native CV-N (**15**) should compete with the CV-N-toxin conjugate and therefore raise the IC_{50} value obtained. This would arise as a consequence of a lesser number of cells being killed due to 'protection' from the native CV-N (**15**) binding to gp120.

The final test will be an ELISA assay to ensure that, unlike the CV-N-PE38 conjugate, the CV-N-halichondrin derivatives will effectively bind to gp120 at the same level as the native protein. The CV-N-PE38 conjugate bound gp120 several orders of magnitude less than the native CV-N (**15**), however, the very much smaller biolinker-halichondrin moiety should not inhibit gp120 binding to this extent.

6.6 Conclusions from Testing

The initial testing of the CV-N-dye conjugates containing the fluorescein and BODIPY fluorescent tags were both essential and informative for several reasons. The first crucial aspect was whether 'blocking' of the amines in CV-N would prevent binding of CV-N to gp120. All the results from the ELISA studies of these conjugates revealed that the derived EC_{50} values remained comparable to that of native CV-N (**15**), no matter how 'loaded' the protein became. It could have been expected that with conjugation of such molecules at distinct sites around the CV-N that the binding affinity for gp120 would be affected to some noticeable extent. The fact that this did not happen was reflective of the unique, and as yet not understood, binding process between CV-N (**15**) and gp120. The XTT assay results from these CV-N-dye conjugates also reflected the lack of interference in the binding to gp120, but this time in a situation that more closely resembled the 'real' virucidal infection process.

The CV-N-biolinker-coumarin conjugates much more closely resembled the final toxin conjugates in that they contained a pendant chain attached to the tetrapeptide biolinker. These conjugates also proved to bind to gp120 just as effectively as the native CV-N (**15**) in the XTT assay. The molecular weight of each coumarin construct per addition was approximately 660 Da which corresponded to the addition of an extra 6% (by weight) to the CV-N (**15**) protein. This was a significant increase in mass particularly if more than one construct was attached per CV-N molecule. However, these additions of mass to the outer surface of the CV-N (**15**) protein, which must extend some distance out from the protein surface, did not inhibit the gp120 binding in any perceivable manner. The lack of fluorescent response from these coumarin conjugates in the gp120-expressing cell assay proved little either way as to whether the biolinker can be cleaved as planned, as there are many factors that can interfere with the fluorescent response from the coumarin.

It was unfortunate that the results from the toxin conjugates themselves could not be included. The testing and the timing of the testing is part of the collaborative effort with

the NCI. The questions left to answer are whether or not these toxin conjugates do selectively target and kill infected, as opposed to uninfected, cells. The other results from this chapter do point towards the toxin conjugates having little trouble binding efficiently to gp120. Based on these results it would be surprising if the conjugates did not bind to gp120 in a comparable manner to native CV-N (**15**). Although the biolinker-toxin section of the conjugate is nearly three times larger than the biolinker-coumarin equivalent, the portion of the structure closest to the CV-N surface (pendant chain-biolinker) is essentially identical. The 'extra' mass comes from the toxin portion of the conjugate which is furthest from the protein surface. The experiments with gp120-expressing infected cells treated with these conjugates is the ultimate test to determine the effectiveness of the toxin conjugates at selectively targeting and destroying HIV-infected cells.

Chapter 7

CONCLUSIONS

7.1 Summary and Conclusions

The aim of this project was to construct CV-N-toxin conjugates that would selectively target HIV-infected cells with subsequent release of the toxin, causing cell death. The early approaches outlined in Chapter 2 led to the choice of two complementary strategies that relied on the activation of toxin constructs by succinimidyl esters and maleimides respectively. The cleavable biolinker provided a suitable method for release of the toxin once internalised in infected cells. The successful development of norhomohalichondrin B amine precipitated its use as the toxin component of the conjugates. The individual components of a construct: protein, activating moiety, biolinker, and toxin had to be incorporated in a coherent strategy that also allowed for the sensitivity of the toxin to harsh chemical conditions.

The native CV-N (**15**) was purified by reverse phase HPLC from a crude mixture obtained from the NCI. This process returned approximately 15% of the mass. Although time consuming, HPLC did effectively separate the CV-N from other proteins present. CV-N-Cys was recombinantly produced, but the presence of a glutathione molecule attached to the cysteine proved troublesome to remove without loss of the biological activity of CV-N. The final derivative of CV-N was produced *via* thiolation of lysine side chain amines. This degree of thiolation could be controlled by halting the reaction at whatever level of conversion was required. Therefore, these CV-N derivatives had varying degrees of 'loading' capacity with the native CV-N (**15**) having the potential for conjugation at six sites, the CV-N-Cys at a single site and the thiolated-CV-N at whatever level in between that was required.

Three separate CV-N-toxin conjugates were ultimately synthesised, all similar in structure, although with slight changes in the coupling process. The succinimidyl ester derived synthetic scheme, although longer than initially planned due to the necessity for lengthening of the chain *via* adipic anhydride, did produce the desired product. The subsequent reaction with CV-N (**15**) was significantly slower than that observed for the model systems incorporating succinimidyl esters. However, the CV-N-toxin product was obtained utilising this methodology and therefore remains a viable strategy.

The synthesis of the maleimido-toxin derivatives utilised a simpler strategy, primarily due to the orthogonal reactivities of the toxin amine and the maleimide moiety. This allowed synthesis of the entire maleimido-caproic-biolinker structure by SPPS with only the carboxyl activation and reaction with the halichondrin amine outstanding. The simple nature of this strategy was a distinct advantage over the succinimidyl counterpart in terms of effort required and exposure of the halichondrin skeleton to stronger (adverse) reaction conditions. This maleimido-toxin conjugate was utilised with both CV-N-Cys and thiolated-CV-N. As noted, the CV-N-Cys provided a derivative with the toxin attached at a defined position. In terms of the viability of CV-N-toxin conjugates as potential therapeutic agents, this homogeneity is of particular importance. However, it will be necessary for future work that CV-N-Cys be made without attached groups such as methyl and glutathione. The thiolated-CV-N reaction was also successful in producing a CV-N-toxin conjugate. This method has the advantage of being able to control the degree of loading of toxin per CV-N molecule by monitoring and halting the thiolation reaction at whatever stage is required, but as yet the site, or most likely sites, of mono-thiolation are unknown.

No matter what modifications were made to CV-N, the initial issue was whether those alterations would cause a decrease in binding affinity for gp120. The results from testing of both the model CV-N-dye and CV-N-biolinker-coumarin conjugates clearly indicated that no discernible inhibition of gp120 binding had taken place. Therefore, it was reasonable to infer that the unique binding properties of CV-N (**15**) to gp120, although not yet precisely understood, did not necessitate interaction between gp120 and at least the

majority of the freely accesible lysine side chains of CV-N (**15**). The extensive testing of these conjugates meant that there was a high degree of confidence that the equivalent toxin constructs would also bind gp120 in the same manner. The final testing of the CV-N-toxin conjugates with infected cells will reveal if the last phase of the strategy, cleavage of the biolinker and release of the toxin, will be successful and a new approach to anti-HIV therapy realised.

Chapter 8

EXPERIMENTAL

8.1 General Methods

Nuclear Magnetic Resonance

All proton detected NMR spectra were recorded on either a Varian Unity 300 or 500 spectrometer at 23^0C, operating at 300 MHz and 500 MHz respectively. Carbon detected NMR spectra were recorded on either a Varian Unity 300 or Varian XL300 spectrometer at 23^0C, operating at 75 MHz. Other NMR experiments described in this thesis *viz* COSY, and the reverse detected HSQC and HMBC experiments were recorded on either the Unity 300 or 500 spectrometer, at 300 MHz or 500 MHz. For each experiment the selection of the particular spectrometer is stated. Chemical shifts in this thesis are described in parts per million (ppm), on the δ scale, and were referenced to the appropriate solvent peaks: $CDCl_3$ referenced to $CHCl_3$ at δ_H 7.26 ppm (^{1}H) and $CHCl_3$ at δ_C 77.0 ppm (^{13}C); CD_3OD referenced to CHD_2OD at δ_H 3.31 ppm (^{1}H) and CD_3OD at δ_C 49.3 ppm (^{13}C). ^{1}H NMR spectra were recorded using an acquisition time (AT) of 2.0 s; ^{13}C NMR spectra were recorded using an AT of 0.878 s. COSY experiments were recorded using an AT of typically 0.215 s and a relaxation delay (D1) of 1.0 s. HSQC experiments with the Pulsed Field Gradient system were run with an AT of typically 0.14 s, a D1 of 1.0 s and $J_{C\text{-}H}$ of 140 Hz. HMBC experiments were recorded with an AT of typically 0.18 s, a relaxation delay of 0.333 s, J=130 Hz and a $^nJ_{CH}$ of 8.3 Hz. HMBC experiments with the Pulsed Field Gradient system were run with an AT of *ca* 0.2 s, a D1 of 1.0 s, J =140 Hz and a $^nJ_{CH}$ of 8.3 Hz.

Mass Spectrometry

Mass spectrometry of some initial samples (stated) was performed on a Kratos MS80 RFA Mass Spectrometer operated at 4,000 V in either fast atom bombardment (FAB) or electron impact (EI) ionisation mode. FAB was performed with an Ion Tech ZNIIFN ion gun using Xe as the reagent gas, operating at 8 kV and 2 mA with an *m*-nitrobenzyl alcohol (NOBA) matrix. Ionisation was performed at 70 eV.

The majority of LCMS samples were analysed with a Waters 2790 HPLC system and a Waters 996 photodiode array (PDA) detector coupled in parallel to a Micromass LCT mass spectrometer equipped with an electrospray ionisation (ESI) probe. Samples were analysed at a probe voltage of 3,200 V at 150^0C with a nebuliser gas flow of 160 L/hr and desolvation gas flow of 520 L/hr with the source temperature at 80^0C. The cone voltage was typically 25 V unless otherwise stated. The solvent flow from a syringe pump in direct injection ESIMS mode (LC disabled) was 20 μL/min.

LCMS of CV-N containing samples used a Zorbax C3 column (150 x 2.1 mm, 300Å, 0.2 mL/min) with a 20%-60% CH_3CN/H_2O (0.5% formic) gradient over 20 minutes unless otherwise stated.

Some samples (stated) were analysed by Dr Lewis Pannell (NIDDK, NIH) on a Hewlett Packard LCMS mass spectrometer.

High Performance Liquid Chromatography

The HPLC work described in this thesis was performed on one of two machines. A Shimadzu LC-4A instrument equipped with a Shimadzu UV Spectrotometric Detector SPD-2AS and Hewlett Packard 3390A integrator was used for preparative reverse phase

work with the stated column, solvent mixture and flow rate. Solvents were degassed using a flow of helium.

Analytical and small scale preparative work was performed on a Shimadzu VP system. The complete setup involved a Shimadzu LC-10AC VP liquid chromatograph coupled to a SIL-10A VP autoinjector, a CTO-10A VP column oven set to 40°C, and a SPD-M10A VP diode array detector. This system was controlled by Shimadzu CLASS-VP (Version 5.023) software. A Shimadzu degasser was utilised for the degassing of all solvents used in this machine.

Columns and flow rates used included a Phenomenex C18 column (250 x 4.6 mm, 60 Å, 5 micron APD, 1 mL/min), a Zorbax C3 column (150 x 2.1 mm, 300 Å, 0.2 mL/min), a Dynamax C4 (260 x 10 mm, 300 Å, 3.5 mL/min) and a Dynamax C18 (270 x 20 mm, 60 Å, 12 mL/min).

The solvents used were either mixtures of acetonitrile (BDH HiperSolv^{TM} 'Far UV' grade) or methanol (BDH HiperSolv^{TM} grade) with water (purified using a MilliQ deionising system).

HPLC purification of CV-N-Cys took place while a guest at the LDDRD, NCI, Frederick, Maryland, USA and used a BioCAD SPRINT Perfusion Chromatography System (PerSeptive Biosystems, Inc.) with a Dynamax C18 column (250 x 10 mm, 300 Å, 2 mL/min).

Bradford Protein Assay

The Bradford protein assay, first reported in 1976[163] is a simple, rapid, and sensitive test for the presence of protein material in a sample. When Coomassie Brilliant Blue is in a dilute acid solution it appears red, but when mixed with protein the dye turns blue (maximum at 595 nm). Not all proteins are susceptible in the assay but CV-N-Cys was

known to provide a positive result. For this particular assay a standard curve was obtained from measuring the response from increasing concentrations of Bovine IgG with the Coomassie Brilliant Blue dye. The crude CV-N-Cys was also tested at two different concentrations. After reading the absorbance at 595 nm for the CV-N-Cys samples, the corrections were made for the dilution factors providing an average value of ~75 mg of protein material in the crude CV-N-Cys sample.

SDS-PAGE

In PAGE, a gel made by polymerising acrylamide and bisacrylamide together forms a network of pores, the sizes of which are defined by the initial acrylamide and bisacrylamide concentrations. When an electric field is applied across the gel, protein migrates according to its charge and the distance that it migrates in the gel in a given time is determined by the pore size, the molecular size of the protein and its net charge. In SDS-PAGE, the anionic detergent sodium dodecyl sulphate (SDS) is used to completely suppress the native charge on the proteins, giving them a large negative coat of detergent molecules. Because SDS interacts with the hydrophobic core of proteins, it causes a rapid and irreversible unfolding that denatures the polypeptide chains which all now have a roughly equivalent charge/mass ratio. For this reason the separation in the gel is now purely dependent on the molecular mass of the protein, providing a useful and high resolution analytical method.

The following general method was used for all SDS-PAGE analyses performed while at the LDDRD, NCI, Frederick, Maryland, USA. 16% Tricine Gels were used which are ideal for low molecular weight proteins such as CV-N-Cys.

The comb was removed from the gel and the wells rinsed with 3x 1 mL of Tricine Running Buffer. The gel cartridge was placed into the XCell II Mini-Cell and buffer poured to cover the gel and one third fill the outer reservoir. To 10 μL of each sample was added 15 μL of Tricine Sample Buffer (see below) which was then vortexed and boiled for 2 min. The samples were loaded into the wells with (usually) a Prestained Marker (see below) in

wells on each end of the gel. The power supply was switched on and the gel run for 75 min. At this time the cartridge was cut open to release the gel and the wells sliced off and removed. The gel was placed in fixative solution (400 mL MeOH, 100 mL CH_3CO_2H, 500 mL de-ionised water) for 30 min then stained with a solution containing CH_3CO_2H (100 mL), Coomassie Blue G-250 (0.25 g) and de-ionised water (900 mL) for 1 hour. The gel was then left overnight in CH_3CO_2H (100 mL) and de-ionised water (900 mL). After washing with water (three times) the gel was placed in Gel-Dry™ drying solution for 20 min. The gel was then sandwiched between two cellophane sheets soaked with Gel-Dry™ and left in a stand overnight to dry.

The Tricine Sample Buffer contained:-

Deionised water (4.0 mL)

0.5M Tris-HCl, pH 6.8 (2.0 mL)

Glycerol (2.4 mL)

10% SDS (1.0 mL)

β-mercaptoethanol (0.2 mL)

0.5% Coomassie G-250 (0.4 mL)

The Prestained Marker contained the following standards:-

Protein	**Molecular Weight (Da)**
Triosephosphate isomerase	26,625
Myoglobin	16,950
α-Lactalbumin	14,437
Aprotinin	6,512
Insulin β chain, oxidised	3,496
Bacitracin	1,423

Any changes made to this protocol are stated.

ELISA Assay

A general outline of the principle and procedure of the gp120-mediated ELISA is discussed in section **6.2.1**. The exact methodology is outlined below.

Each well of a 96-well plate was incubated with 100 µL/well of a 1 µg/mL solution of gp120 (8.3 nM) in PBS for 2 hours at room temperature. Plates were then washed with 100 µL of wash buffer (PBS (without Ca and Mg), 0.01% Tween 20). Incubation with 200 µL of 2% BSA in PBS (sonicated) overnight at 4^0C blocked any unused space in the wells. Plates were then washed twice with 100 µL of wash buffer. The appropriate wells (shown in **Figure 6.2**) were incubated with 100 µL/well of variable concentrations of either native CV-N (**15**) or CV-N conjugates for 1 hour at room temperature. The plate was again washed 3 times with 100 µL of wash buffer and further incubated with 100 µL/well of anti-Rabbit-anti-CV-N Ab for 1 hour at room temperature. After a further 3 washes the plate was incubated with 100 µl/well of anti-Rabbit IgG-Ab-AP for 1 hour at room temperature. The plates were then washed 3 times and incubated with 100 µL of substrate buffer (10% diethanolamine, 10 mM $MgCl_2$, 4 mg/mL *p*-nitrophenyl phosphate (pNPP) for approximately 15 min – dependant on visual colour change. The final step was to add 100 µL of EDTA (0.5 M, pH 8.0) to each well (without washing of the buffer). The plates were then stored at 4^0C until ready to read at 405 nm in a plate reader.

XTT Assay

After adding each sample for testing to a single column of a 96 well plate at a known concentration the rest of the procedure was primarily automated as described in section **6.2.2**. For further detailed information see Weislow *et al.*[162] and Gulakowski *et al.*[164]

Column Chromatography

All column chromatography was performed using glass columns of stated dimensions. Solvents were all of commercial grade, distilled once in glass distillation apparatus, except MeOH, which was distilled twice. 'Flash' columns were run under N_2 gas (oxygen free) pressure (0.5 kPa).

Silica flash chromatography was performed on Merck silica gel 60 (230-400 mesh). Reverse phase C18 chromatography of halichondrin containing compounds used Bakerbond spe^{TM} Octadecyl (C18) disposable extraction columns (40 μm APD, 60 Å) of either 100 mg or 500 mg size. Chromatography of CV-N conjugates used Bakerbond C4, Pharmacia G25 or G100 in appropriate solvents (stated).

Kaiser Test

For SPPS of peptides the Kaiser test was used to determine the presence (or not) of free amines. Several beads from the resin synthesis were removed and washed with 3x 1 mL EtOH in a small glass vial. The following three stock solutions were made and 3 drops from each added to the beads ensuring all beads were completely covered in solution.

1) 2.5 g of ninhydrin in 50 mL of EtOH
2) 40 g phenol in 10 mL of EtOH
3) 0.001 M KCN (6.5 g in 100 mL water) and 12 mL of pyridine

The vial was capped and placed in an oven (100^0C) for approximately 5 min. A positive result was indicated by the beads turning blue while a negative result was shown by no change in colour (beads remain yellow).

Solvents

All Technical grade solvents were distilled prior to use. MeOH was distilled twice. Dry solvents were obtained using the following standard methods. MeOH was distilled from magnesium metal and iodine and stored over activated molecular sieves (4 Å). Pyridine was refluxed over CaH_2 before distillation. Tetrahydrofuran (THF) was refluxed over sodium metal and benzophenone before distillation directly prior to use. Dichloromethane (DCM) was refluxed over calcium hydride before distillation. DMF was dried by treating twice overnight with 4 Å molecular sieves, followed by storage over 4 Å sieves. Isopropyl alcohol (IPA) was refluxed over calcium hydride for 3 hours before distillation and stored over 4 Å sieves.

8.2 Work Described in Chapter 2

2.6.2.1 Reductive Amination of Norhomohalichondrin B Aldehyde

To a solution of **29** (1.0 mg, $9.17x10^{-4}$ mmole) in dry MeOH (200 μL) was added ammonium acetate (7.06 mg, 0.0917 mmole) and $NaCNBH_3$ (0.0403 mg, $6.42x10^{-4}$) with stirring under argon at room temperature. TLC analysis (SiO_2, Merck, 10% MeOH/DCM) after 4 hours showed formation of a ninhydrin and phosphomolybdic acid (PMA) active spot at R_f=0.2 compared to the R_f =0.8 of the PMA active **29**. After 18 hours all signs of **29** had disappeared (TLC) and the reaction mixture was loaded in 50 μL aliquots onto a C18 cartridge (100 mg) pre-filled with water (~5 column volumes) above the packing material. After all 4 aliquots had been diluted and rinsed through the cartridge in this way the column was washed with water (20 column volumes). Washing with MeOH (10 column volumes) eluted **28** in quantitative yield after drying. The ^{1}H NMR (Unity 500) spectrum of **28** was directly compared with that of the starting aldehyde (**29**) and amine synthesised previously *via* a different method.[130] The disappearance of the aldehyde resonance at δ 9.70 ppm and the movement of resonances for H53 and H50 along with the correlation to the previously derived amine indicated successful product formation. Confirmation by ESIMS analysis showed a mass of 1092.48 Da. $C_{60}H_{86}NO_{17}$ (MH^+) required 1092.58.

8.3 Work Described in Chapter 3

Work in this chapter from sections' **3.3.1**, **3.3.2**, **3.3.3.1**, **3.3.3.2**, **3.3.3.3** was carried out while a guest at the LDDRD, NCI, Frederick, Maryland, USA.

3.2 Native CV-N Purification

Successive analytical HPLC injections (Phenomenex C18 column) of the crude CV-N (1.712 g) showed that the best method was a 20% to 50% CH_3CN/H_2O (0.05% TFA) gradient over 15 min with monitoring at 210 nm. The CV-N (**15**) eluted at 13.1 min, confirmed by ESIMS of the collected peak showing a deconvoluted mass of 11,009 Da. Preparative HPLC on the Shimadzu LC-4A utilised a 20% to 45% gradient over 25 min using the same solvents. The average injections were 200 µL of a 25 mg/mL solution of crude CV-N onto a Dynamax C4 column with a flow rate of 4 mL/min. The CV-N (**15**) peak (22 min) was collected and after multiple injections dried under vacuum. The mass obtained for the CV-N (**15**) peak was 12% of the total mass injected. Re-injection of the collected peak on the analytical HPLC system confirmed the identity of the CV-N (**15**) with an identical retention time to a CV-N (**15**) standard and distinctive UV chromophore characteristic of protein material. A sample sent to Dr Lewis Pannell (NIDDK, NIH) for LCMS confirmed the mass as 11,009 Da corresponding to $C_{470}H_{743}N_{133}O_{164}S_4$.

3.3.1 Recombinant Production of CV-N-Cys

Cultures of transformed *E. coli* containing the plasmid (pET26b(+)) with coding for CV-N-Cys were grown in Super Broth (Quality Biological Inc., Gaithersberg, USA) with kanamycin (30 µg/mL), 0.5% glucose, and 1.6 mM $MgSO_4$ at 37^0C overnight. 20 mL of this broth was added to each of eight 2L flasks containing 500 mL of Super Broth and incubated for a further 4 hours at 37^0C. When the OD of the cultures (600 nm) had reached

a value of ~4, IPTG was added to a final concentration of 1 mM for induction of gene expression. Cells were harvested after 2 hours of incubation at 37^0C through centrifugation at 6,500 g for 15 min at 4^0C. The pelleted cells were resuspended in Tris-HCl (30 mM, pH 8.0), 20% sucrose, and EDTA (1 mM), incubated for 15 min at room temperature with stirring and centrifuged at 10,000 g for 10 min at 4^0C. The resulting supernatant (periplasmic fraction) was freeze dried.

3.3.2 Purification and Analysis of CV-N-Cys

The freeze dried solid (from above) was resuspended in water and chromatographed on a pad of reverse phase C4 in a 250 mL sintered glass funnel. The sample was loaded in water and progressively washed with 100% water, 2:1 water/MeOH, 1:2 water/MeOH and 100% MeOH. The MeOH containing fractions were combined and dried under vacuum. A Bradford Protein Assay estimated the protein content to be 75 mg at this stage and SDS-PAGE analysis showed, as the primary spot, a protein of approximately 11 kDa in size. Purification was achieved firstly by reverse phase C18 preparative HPLC (BioCAD SPRINT Perfusion Chromatography System) running a 20% to 50% CH_3CN/H_2O (0.05% TFA) gradient over 30 min with monitoring at 210 nm and a flow rate of 2 mL/min. The CV-N-Cys eluted at 22 min, was collected over multiple injections and dried under vacuum to yield 43 mg. A further SDS-PAGE established that the sample had undergone significant purification with a single spot at approximate molecular weight 11 kDa. LCMS analysis by Dr Lewis Pannell (NIDDK, NIH) showed the molecular mass to be 11,417 Da, 305 Da higher than would be expected for the CV-N-Cys protein. The extra mass was determined to be a glutathione molecule by comparison to previous samples analysed from the LDDRD that had contained smaller amounts of this product.

3.3.3.1 Dialysis Treatment of CV-N-Cys-Glutathione

The initial dialysis buffering solution (1 L) contained potassium phosphate buffer (20 mL, 100 mM), EDTA (2 mM, pH 7.5) and DTT (100 μL, 1 M). To **30** (250 μL of 4 mg/mL) was added 2.25 mL of dialysis buffer and the solution injected into the inner membrane of the dialysis cartridge. The cartridge was left submerged in the larger volume at 4^0C with gentle stirring for 6 hours at which time the outer solution was refreshed with 1 L of the same buffering solution. After a further 12 hours, the outer solution was changed to 1 L of buffer (no DTT) and left stirring for 6 hours with a final change to 1 L of buffer (no DTT) which was left stirring at 4^0C for 12 hours. The inner solution was removed from the cartridge and chromatographed on a C4 column (30 x 10 mm) and washed with water (20 column volumes). The protein was eluted with MeOH (10 column volumes) and dried under vacuum. An initial SDS-PAGE analysis under standard conditions revealed protein corresponding to approximately the monomer of CV-N-Cys but a further SDS-PAGE without boiling of the sample showed protein corresponding to the dimer. This was confirmed by LCMS analysis (Dr Lewis Pannell, NIDDK, NIH). Mass of 22,262.79 Da (dimer of CV-N-Cys plus K^+). An XTT assay showed an EC_{50} 12x less than that of native CV-N (**15**).

3.3.3.2 Repeat of Dialysis of CV-N-Cys-Glutathione

The initial dialysis conditions were identical to those above. After stirring for 24 hours the inner solution was removed from the cartridge and diluted into water (500 mL) with slow stirring. After 1 hour the solution was progressively loaded into two 80 mL Centriprep filters (3,000 MW cut-off) which were centrifuged at 3,000 g for 20 min. After the 500 mL had been centrifuged through the filter the remaining small volume (~5 mL in each filter) was further washed and centrifuged with water (3x 70 mL). The filter was then reversed and centrifuged at 250 g for 1 min. The samples from both filters were combined and freeze dried. SDS-PAGE analysis revealed monomer with pre-boiling of the sample, and dimer without boiling. LCMS (Dr Lewis Pannell, NIDDK, NIH) confirmed the presence

of dimer of mass 22,223.26 Da. XTT analysis gave an EC_{50} 60 times greater than native CV-N (**15**).

3.3.3.3 Chemical Removal of Glutathione from CV-N-Cys

To **30** (10 mg, $8.76x10^{-4}$ mmole) in 250 μL of water was added guanidine HCl (750 μL, 8 M) with gentle vortexing and addition of Tris HCl (33 μL of 3 M, pH 8.45) and DTT (20 μL of 0.5 M). The sample was flushed with N_2, sealed, left in the dark for 2 hours and then chromatographed on a G100 column (400 x 25 mm) eluting with PBS at a flow rate of 2.1 mL/min with monitoring at 280 nm. The protein was collected between 32 and 52 min and concentrated on Centriprep filters (as described for the repeat dialysis experiment above) and freeze dried. SDS-PAGE analysis with pre-boiling showed monomer, and without boiling an approximate ratio of 1:2 monomer:dimer. LCMS analysis (Dr Lewis Pannell, NIDDK, NIH) showed only monomer of correct mass 11,112.04 Da. However, an XTT assay showed an EC_{50} 668 times greater than native CV-N (**15**).

3.4.1 Lysozyme Model with Iminothiolane

Lysozyme (1 mg, $6.99x10^{-5}$ mmole) was dissolved in 1 mL of degassed (helium) reaction buffer (0.1 M sodium borate, 0.001 M EDTA, 0.15 M NaCl) with addition of iminothiolane (**27**, 96 μL of 1 mg/mL, $6.99x10^{-4}$ mmole). Aliquots (5 μL) were removed after 5, 25, 45, 65 and 90 min, diluted 30 fold with water and injected into the ESIMS probe with 5 μL of a 10% formic acid solution. Progressive formation of increasing numbers of thiolated amines was noted (see **Figure 3.4** for all spectra) and after 90 min the solution was further reacted (section **5.2.1**).

3.4.2.1 Preliminary Reaction to Test Effectiveness of Thiolation

To CV-N (**15**, 200 μg, $1.82x10^{-5}$ mmole) in 200 μL of a 0.9% saline solution was added 200 μL of reaction buffer (0.1 M sodium borate, 0.001 M EDTA, 0.15 M NaCl) with degassing (helium) followed by addition of iminothiolane (**27**, 25 μL of 1 mg/mL, $1.82x10^{-4}$ mmole). Aliquots (10 μL) were removed after 5 and 25 min, diluted 30 fold with water and injected into the ESIMS probe with 10 μL of a 10% formic acid solution. The reaction was proceeding slowly so further iminothiolane (**27**, 50 μL of 1 mg/mL, $3.64x10^{-4}$ mmole) was added with ESIMS analysis at a total reaction time of 30, 45 and 60 min. The conversion rate from amines to thiols was still slow, so further iminothiolane (**27**, 50 μL of 1 mg/mL, $3.64x10^{-4}$ mmole) was added to the reaction mixture. A final ESIMS analysis at total reaction time of 80 min (see **Figure 3.5** for all spectra) showed sufficient thiolation to proceed with the next reaction (section **5.2.2**).

3.4.2.2 Second Reaction with Modified Conditions

To CV-N (**15**, 200 μg, $1.82x10^{-5}$ mmole) in 200 μL of a 0.9% saline solution was added 200 μL of reaction buffer (0.1 M sodium borate, 0.001 M EDTA, 0.15 M NaCl) with degassing (helium) followed by addition of iminothiolane (**27**, 25 μL of 5 mg/mL, $9.10x10^{-4}$ mmole). Aliquots (5 μL) were removed after 60 and 105 min, diluted four fold with water and injected into the ESIMS probe with 15 μL of a 10% formic acid solution. The reaction mixture showed formation of significant amounts of thiolated-CV-N and was chromatographed on a G25 column (400 x 10 mm) eluting with water at a flow rate of 0.9 mL/min and monitoring at 276 nm. The separated protein fraction was further reacted (section **5.3.2**).

3.4.2.3 Optimised Reaction of CV-N Plus Iminothiolane

To CV-N (**15**, 200 μg, $1.82x10^{-5}$ mmole) in 200 μL of a 0.9% saline solution was added 200 μL of reaction buffer (0.1 M sodium borate, 0.001 M EDTA, 0.15 M NaCl) with degassing (helium) followed by addition of iminothiolane (**27**, 50 μL of 5 mg/mL, $1.82x10^{-3}$ mmole). ESIMS injections at 5 min and 25 min showed formation of principally singly thiolated-CV-N. The solution was left for 30 min and then chromatographed on a G25 column (400 x 10 mm) eluting with water at a flow rate of 0.9 mL/min and monitoring at 276 nm. The protein eluted between 10 and 14 min. ESIMS analysis confirmed the formation of mostly singly thiolated-CV-N (see **Figure 3.6** for spectrum). Half of this solution from the G25 column was reacted (section **5.5.2**).

8.4 Work Described in Chapter 4

Work in this chapter from sections' **4.2.2.1**, **4.2.2.2**, **4.2.3.1** and **4.2.3.2** was carried out while a guest at the LDDRD, NCI, Frederick, Maryland, USA.

4.2.2.1 Reactions at pH 8.2

Four reactions were performed on separate aliquots of CV-N (**15**, 200 μg, 1.82×10^{-5} mmole) dissolved in $NaHCO_3$ buffer (200 μL, 0.1 M). The 5-SFX dye (1 mg) was dissolved in DMF (100 μL) and aliquots (2 equiv, 2.14 μL; 6 equiv, 6.41 μL; 2 x 25 equiv, 26.7 μL) added to the CV-N solutions. The first three reactions were left at room temperature for 15 min. The fourth reaction, with 25 equiv. was allowed to react for 60 min. After completion, the samples were chromatographed on a G25 column (140 x 10 mm) eluting PBS. Two coloured bands (light orange, bright yellow) for each column were observed and subsequently collected. The first bands eluting were the protein containing fractions which were submitted for LCMS analysis (Dr Lewis Pannell, NIDDK, NIH). The results shown in **Table 4.1** and **Figure 4.1**.

4.2.2.2 Reactions at pH 7.2

The same reactions as above were carried out with the only differences being the buffer (0.1 M PBS, pH 7.2) and the less intense colour of the eluted protein bands from the G25 column. The results from LCMS analysis (Dr Lewis Pannell, NIDDK, NIH) shown in **Table 4.2** and **Figure 4.3**.

4.2.3.1 Reactions at pH 8.2

Reactions were as described in **4.2.2.1** above, but with the following quantities of BODIPY FLX-SE (1 mg in 100 μL of DMF) used in place of 5-SFX; (2 equiv, 1.83 μL; 6 equiv, 5.48 μL; 2 x 25 equiv, 22.8 μL). The colours of the eluted bands were orange and light green. The protein containing bands (orange) were also analysed by LCMS (Dr Lewis Pannell, NIDDK, NIH) with results shown in **Table 4.3** and **Figure 4.4**.

4.2.3.2 Reactions at pH 7.2

The same reactions as above were carried out with the only differences being the buffer (0.1 M PBS, pH 7.2) and the less intense colour of the eluted protein bands from the G25 column. The results from LCMS (Dr Lewis Pannell, NIDDK, NIH) shown in **Table 4.4** and **Figure 4.5**.

4.3.1.2 Synthesis of Fmoc-Gly-Phe-Leu-Gly-OH

The SPPS of **49** was achieved by standard solid phase techniques using a glass apparatus with a frit (porosity 3) allowing bubbling of the resin from below through flow of N_2, and subsequent vacuum removal of the solution after reaction at each stage. Fmoc-Gly-Wang resin (3 g, 0.897 meq/g) was bubbled for 30 min in dry DMF, drained, and the Fmoc cleaved with bubbling of a 20% piperidine/DMF solution (30 mL) for 10 min. After washing with 3 x 15 mL of DMF and 3 x 20 mL IPA, several beads were removed and analysed using the Kaiser test (see General Methods, section **8.1**). A positive result allowed washing with 3 x 15 mL DMF and changing of the flask containing all the washings (to remove the presence of the piperidine). A DMF solution (~15 mL) containing Fmoc-Leu (1.902 g, 5.382 mmole), HBTU (2.041 g, 5.382 mmole), HOBt (0.727 g, 5.382 mmole), and DIPEA (1.875 mL, 10.764 mmole) was bubbled with the resin for 60 min. Washing with 3 x 15 mL DMF and 3 x 20 mL IPA was followed by removal of several

more beads for a Kaiser test, this time negative, indicating that all amines had reacted. The process of coupling and deprotection was repeated for addition of Fmoc-Phe (2.085 g, 5.382 mmole) and Fmoc-Gly (1.600g, 5.382 mmole) with all other reagents the same as above. The tetrapeptide was left to dry overnight inside the glass vessel in the presence of P_2O_5 under vacuum. A mixture of 45 mL of 95% TFA, 2.5% TES and 2.5% water was bubbled with the resin for 20 min, drained, and then rinsed with 2 x 20 mL TFA. The combined acidic solution was alternately concentrated under vacuum and repeatedly precipitated with water to remove TFA from solution. The solution was precipitated one last time and filtered over a sintered glass funnel (porosity 4), and dried to give 1.17 g of **49** (71% yield). HRFABMS showed MNa^+ 637.2639 ($C_{34}H_{38}O_7N_4Na$ required 637.2657). ^1H (Unity 300) and ^{13}C NMR (XL300) spectral data shown in **Table 4.5**.

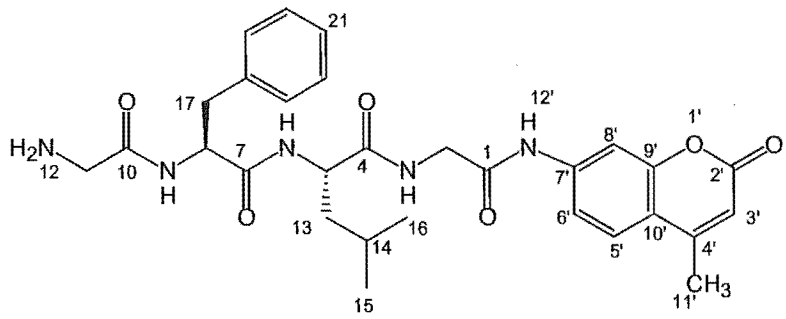
4.3.2.1 Synthesis of Fmoc-Gly-Phe-Leu-Gly-Coumarin

To **49** (30 mg, 0.0488 mmole) in dry pyridine was added 7-amino-4-methylcoumarin (**50**, 9.4 mg, 0.0537 mmole) and $POCl_3$ (5.0 µL, 0.0537 mmole) with stirring for 1 hour. After removal of solvent, the residue was dissolved in MeOH and an analytical HPLC (Phenomenex C18 column) performed (80% MeOH/H_2O (0.05% TFA)) with elution of the product **51** at a retention time of 8.24 min compared to **49** at 6.38 min. Preparative HPLC on the Shimadzu LC-4A (Dynamax C18) of **51** with the same isocratic solvent mixture gave a retention time of 27 min for the product which was collected, dried and weighed (25.3 mg, 67%). HRESIMS gave MH^+ 772.3341 ($C_{44}H_{46}N_5O_8$ required 772.3346). ^1H (Unity 300) and ^{13}C NMR (XL 300) spectral data shown in **Table 4.6**.

4.3.2.3 Fmoc Removal to form H_2N-Gly-Phe-Leu-Gly-Coumarin

51 (25.3 mg, 0.0328 mmole) was dissolved in 20% piperidine/CH_3CN (5 mL) and left for 30 min and then evaporated to dryness. The residue was dissolved in DCM and chromatographed on a silica column (60 x 5 mm) and washed initially with DCM (5 mL).

Elution with EtOAc (10 mL) yielded the Fmoc-piperidine adduct (ESIMS), followed by elution with MeOH (10 mL) which yielded the product (**53**, 16.8 mg, 93%). The ^{1}H NMR (Unity 300) spectrum in CD_3OD was assigned through comparison with **51**.



H_2N-Gly-Phe-Leu-Gly-coumarin (**53**)

^{1}H NMR (CD_3OD) δ 0.91 (d, *J*=6.1 Hz, $(H15)_3$ or $(H16)_3$), δ 0.95 (d, *J*= 6.1 Hz, $(H16)_3$ or $(H15)_3$), δ 1.52-1.71 (m, $(H13)_2$), δ 1.52-1.71 (m, H14), δ 2.44 (s, $(H11')_3$), δ 2.96 (dd, *J*=8.1 and 14.3 Hz, H17a), δ 3.16 (dd, *J*=5.7 and 14.3 Hz, H17b), δ 3.47 (m, $(H11)_2$), δ 3.88 (d, *J*=17.1 Hz, H2a), δ 4.14 (d, *J*=17.1 Hz, H2b), δ 4.31 (t, *J*=6.3 Hz, H5), δ 4.67 (t, *J*=5.8 Hz, H8), δ 6.21 (s, H3'), δ 7.15-7.23 (m, 2x H19), δ 7.15-7.23 (m, 2x H20), δ 7.15-7.23 (m, H21), δ 7.66 (s, H5'), δ 7.66 (s, H6'), δ 7.76 (s, H8') amides were not observed

4.3.3 Succinimide Activation of H_2N-Gly-Phe-Leu-Gly-Coumarin

This reaction was performed many times with a typical procedure for the succinimide activation shown below. To H_2N-Gly-Phe-Leu-Gly-coumarin (**53**) in dry CH_3CN was added DSC (**54**, 2 equiv.) and pyridine (2 equiv.). [Other reaction conditions attempted included a change of base (triethylamine, DIPEA), solvent (DMF, DCM) and performing the reaction under argon.] The reaction was monitored by TLC analysis (silica, 20% MeOH, EtOAc) and after 4 hours an increase in the R_f of the starting material from 0.2 to 0.8 (UV, vanillin) was observed. The reaction mixture was dried, with a crude ^{1}H NMR spectrum typically showing the presence of a singlet at 2.8 ppm representing the four protons of the succinimide ring. An ESIMS analysis at this stage in general showed the

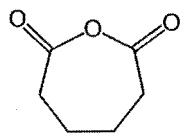
correct mass for the product. Purification of the crude sample was required but whatever method was used, such as column chromatography (silica, diol) or HPLC (C18), the product degraded forming a mixture of compounds with the dominant species having a mass 26 Da higher than **53**. The ^{1}H NMR spectrum of the 'purified' sample was generally a mixture, but a downfield shift in the α-proton of phenylalanine by 0.50 ppm and the β-protons by 0.16 ppm respectively was observable. A shift in the *N*-terminal glycine α-protons of 0.22 ppm also suggested a change in the structure. Not enough material was available of sufficient purity to characterise by NMR spectroscopy although a structure was postulated (**Figure 4.11**).

4.3.3.1 Chain Extension Reaction with Succininc Anhydride

A sample of **53** (3.7 mg, 6.73×10^{-3} mmole) was reacted with succinic anhydride (0.81 mg, 8.08×10^{-3} mmole) with DMAP (0.08 mg, 6.73×10^{-4} mmole) in dry pyridine (1 mL). After 24 hours the sample was dried and purified by preparative HPLC (Shimadzu LC-4A, Dynamax C18, 40% CH_3CN/H_2O (0.05% TFA), 3.5 mL/min, 320 nm). An ESIMS analysis showed the correct mass (650 Da, MH^+) for succinic-Gly-Phe-Leu-Gly-coumarin and the product was immediately reacted with DSC (2 equiv.) and pyridine (2 equiv.) in CH_3CN. LCMS analysis (Phenomenex C18, 40% CH_3CN/H_2O (0.5% formic), 320 nm) showed formation of the product with the correct mass although a secondary peak also formed which represented a mass 26 Da higher than the starting material, succinic-Gly-Phe-Leu-Gly-coumarin. Purification by preparative HPLC (Shimadzu LC-4A, Dynamax C18, 40% CH_3CN/H_2O (0.05% TFA), 320 nm) produced primarily the compound of mass 26 Da higher than the starting material. This was again unacceptable and this strategy was modified (section **4.3.4.2**).

4.3.4.1 Adipic Anhydride Synthesis

Adipic acid (5 gm, 0.034 mole) was dissolved in acetic anhydride (9.7 mL, 0.102 mole) and refluxed for 2 hours at 160^{0}C under argon. The excess acetic anhydride and any acetic acid formed were distilled off at normal pressure. The liquid remaining solidified at room temperature and was distilled on a Kugelrohr (Büchi GKR-50) at approximately 180^{0}C under vacuum. LREIMS on the Kratos MS80 found 128 Da ($C_6H_8O_3$).

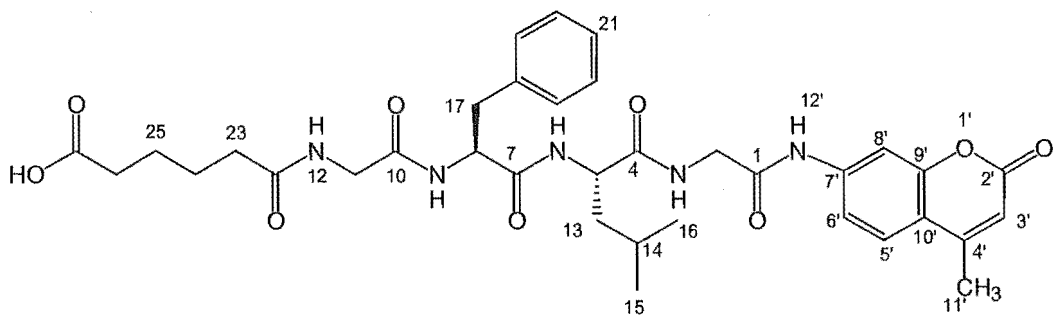


adipic anhydride (**55**)

^{1}H NMR ($CDCl_3$) δ 1.73 (m, 2x CH_2), δ 2.50 (m, 2x CH_2)

4.3.4.2 Adipic Anhydride with H_2N-Gly-Phe-Leu-Gly-Coumarin

To **53** (5.6 mg, 0.0102 mmole) in CH_3CN was added adipic anhydride (1.44 mg, 0.0112 mmole) with stirring for 14 hours. After removal of solvent the sample was purified by preparative HPLC (Shimadzu LC-4A, Dynamax C18, 50% CH_3CN/H_2O (0.05% TFA), 210 nm) with a retention time of 7.0 min for the starting material and 11.6 min for the product (**56**, 4.7 mg, 68%). ESIMS analysis gave 678.41 (MH^+) ($C_{35}H_{44}N_5O_9$ required 678.31). The ^{1}H NMR was recorded on the Unity 500 spectrometer.



adipic-Gly-Phe-Leu-Gly-coumarin (**56**)

^{1}H NMR (CD_3OD) δ 0.92 (d, *J*=5.9 Hz, $(H15)_3$ or $(H16)_3$), δ 0.96 (d, *J*= 5.9 Hz, $(H16)_3$ or $(H15)_3$), δ 1.49 (m, H14), δ 1.62-1.72 (m, $(H13)_2$), δ 1.65 (m, $(H24)_2$), δ 1.65 (m, $(H25)_2$), δ 2.25 (m, $(H23)_2$), δ 2.26 (m, $(H26)_2$), δ 2.46 (s, $(H11')_3$), δ 2.99 (m, H17a), δ 3.16 (m, H17b), δ 3.76 (d, *J*=15.9 Hz, H11a), δ 3.83 (d, *J*=15.9 Hz, H11b), δ 3.91 (dd, *J*=6.3 and 17.1 Hz, H2a), δ 4.14 (dd, *J*=6.3 and 17.1 Hz, H2b), δ 4.29 (m, H5), δ 4.62 (m, H8), δ 6.24 (s, H3'), δ 7.20 (m, 2x H19), δ 7.20 (m, 2x H20), δ 7.20 (m, H21), δ 7.69 (s, H5'), δ 7.71 (s, H6'), δ 7.79 (s, H8'), amide protons were not observed.

4.3.5 Activation of Adipic-Gly-Phe-Leu-Gly-Coumarin

To **56** (2.5 mg, $3.69x10^{-3}$ mmole) in DMF (500 μL) was added DCC (0.76 mg, $3.69x10^{-3}$ mmole) and NHS (0.45 mg, $3.69x10^{-3}$ mmole). After 12 hours, preparative HPLC (Shimadzu LC-4A, Dynamax C18, 45% CH_3CN/H_2O (0.05% TFA), 324 nm) separated the product with a retention time of 14.1 min. (**57**, 1.81 mg, 63%). ESIMS analysis found 775.50 (MH^+) ($C_{39}H_{47}N_6O_{11}$ required 775.32).

4.3.6 Reaction of CV-N with 57

Dual reactions were carried out on two samples of CV-N (200 μg, $1.82x10^{-5}$ mmole) in 0.9% saline with $NaHCO_3$ (1.68 mg, 0.02 mmole, pH 8.2). To these was added **57** (0.352 mg, $4.54x10^{-4}$ mmole) in DMF (50 μL). The first sample was left for 15 min, chromatographed on a G25 column (400 x 10 mm) eluting with water at a flow rate of 0.9 mL/min with monitoring at 276 nm. The protein-containing fraction that eluted first was collected and analysed by LCMS (Zorbax C3, 280 nm). LCMS data shown in **Figure 4.16** and **Figure 4.17**. The second sample had a reaction time of 60 min with an identical analysis as above. LCMS data shown in **Figure 4.18** and **Figure 4.19**.

4.4.1 Synthesis of Fmoc-Gly-Phe-Leu-Gly-Succinimidyl Ester

Fmoc-Gly-Phe-Leu-Gly-OH (**49**, 1.12 mg, 1.82×10^{-3} mmole) was dissolved in dry THF (2 mL) with DCC (0.75 mg, 3.64×10^{-3} mmole) and NHS (0.42 mg, 3.64×10^{-3} mmole), flushed with argon, sealed, and left stirring at room temperature. Analytical HPLC analyses (Phenomenex C18 column, 60% CH_3CN/H_2O (0.05% TFA), 210 nm) showed progressive formation of the succinimidyl ester at a retention time of 9.2 min. After 14 hours, no sign of the starting material was visible by HPLC and the solution containing **60** was ready for addition of **28**.

4.4.2 Synthesis of Fmoc-Gly-Phe-Leu-Gly-Norhomohalichondrin B

To the solution containing the activated ester **60**, was added norhomohalichondrin B amine (**28**, 1.0 mg, 9.16×10^{-4} mmole) in dry THF. Aliquots (1 μL) were removed for ESIMS analysis at various time intervals over 48 hours until complete disappearance of the halichondrin amine and formation of peaks (singly and doubly charged) representing the product (**61**). Preparative HPLC on the Shimadzu VP system (Phenomenex C18 column, 85% CH_3CN/H_2O, 210 nm) separated the product (**61**, retention time 8.4 min) which was collected and dried (0.80 mg, 52%). HRESIMS gave $(MH_2)^{2+}$ 844.9344 ($(C_{94}H_{123}N_5O_{23})/2$ required 844.9304).

4.4.3 Fmoc Removal from 61

61 was dissolved in 2% piperidine/CH_3CN for 30 min and the solvent and piperidine removed under vacuum. HRESIMS gave $(MHNa)^{2+}$ 744.8842 ($(C_{79}H_{112}N_5O_{21}Na)/2$ required 744.8874).

4.4.4 Adipic Anhydride Reaction with 62

To the crude reaction mixture of **62** dissolved in CH_3CN was added adipic anhydride (0.067 mg, $5.21x10^{-4}$ mmole) under argon. The reaction was monitored by ESIMS analysis with completion of the reaction by 26 hours. After drying under vacuum the sample was dissolved in minimal MeOH and chromatographed on a C18 cartridge (100 mg) pre-filled with water (~5 column volumes) above the packing material. After the sample had been diluted and rinsed through the cartridge in this way the column was washed with water (20 column volumes). Washing with MeOH (10 column volumes) eluted **63** which was dried under vacuum (0.62 mg, 82% from **61**). HRESIMS analysis gave $(MNa_2)^{2+}$ 819.9035 ($(C_{85}H_{119}N_5O_{24}Na_2)/2$ required 819.9020).

4.4.5 Activation of Adipic-Gly-Phe-Leu-Gly-Norhomohalichondrin B

To **63** (0.62 mg, $3.89x10^{-4}$ mmole) in dry CH_3CN (0.6 mL) was added DCC (0.387 mg, $1.88x10^{-3}$ mmole) and NHS (0.216 mg, $1.88x10^{-3}$ mmole). An ESIMS injection after 48 hours showed complete conversion to the succinimidyl ester (**64**). Purification by HPLC on the Shimadzu VP system (Phenomenex C18 column, 65% CH_3CN/H_2O, 210 nm) separated the product (**64**) which was collected and dried (0.42 mg, 52%). ESIMS analysis gave $(MNa_2)^{2+}$ 868.82 ($(C_{89}H_{122}N_6O_{26}Na_2)/2$ required 868.41).

4.4.6 Reaction of 64 with CV-N

CV-N (**15**, 50 μg, $4.54x10^{-6}$ mmole) was dissolved in $NaHCO_3$ buffer (150 μL, 0.1 M) and **64** (0.153 mg, $9.08x10^{-5}$ mmole) added in CH_3CN (50 μL). LCMS analysis (Zorbax C3) after 15 hours showed product formation but with CV-N (**15**) still present. A further 20 equiv. of **64** ($9.08x10^{-5}$ mmole) was added. LCMS analysis after a further 15 hours showed an increase in the relative size of the peak for the product. LCMS data shown in **Figure 4.25**. The reaction mixture was separated by HPLC using the same gradient profile and conditions as for LCMS, but with a change in the acid to 0.001% formic acid.

8.5 Work Described in Chapter 5

5.2.1 Thiolated-Lysozyme Reaction with a Maleimide

To the thiolated-lysozyme reaction mixture from section **3.4.1** was added maleimidobutyric acid (**65**, 0.128 mg, 6.99×10^{-4} mmole) in reaction buffer (50 μL, 0.1 M sodium borate, 0.001 M EDTA, 0.15 M NaCl). ESIMS analysis after 5 min with the addition of formic acid to the analysis sample allowed observation of a significant concentration of product. ESIMS data shown in **Figure 5.2**.

5.2.2 Thiolated-CV-N Reaction with F-150

To the thiolated-CV-N formed in section **3.4.2.1** was added F-150 (**66**, 0.00776 mg, 1.82×10^{-5} mmole) in DMF (10 μL). After 1 hour the reaction mixture was chromatographed on a G25 column (400 x 10 mm) eluting with water at a flow rate of 0.9 mL/min with monitoring at 276 nm. The protein containing fraction eluted at 6.7 min with the remainder of the lower molecular weight compounds eluting after 9 min. ESIMS analysis of the protein fraction revealed significant addition of F-150 (**66**) to CV-N (**15**). ESIMS data shown in **Figure 5.4**.

5.3.1 Synthesis of Maleimido-Caproic-Gly-Phe-Leu-Gly-Coumarin

To H_2N-Gly-Phe-Leu-Gly-coumarin (**53**, 2.0 mg, 3.64×10^{-3} mmole) in DCM (0.5 mL, 3 drops of DMF for solubility) was added maleimido-caproic acid (**67**, 3.64×10^{-3} mmole) and DCC (1.12 mg, 5.46×10^{-3} mmole). TLC analysis (silica, 20% MeOH/EtOAc) after 20 hours showed formation of a fluorescent spot (R_f=0.6, vanillin) more non-polar than the baseline starting material (**53**). After removal of solvent under vacuum the product was

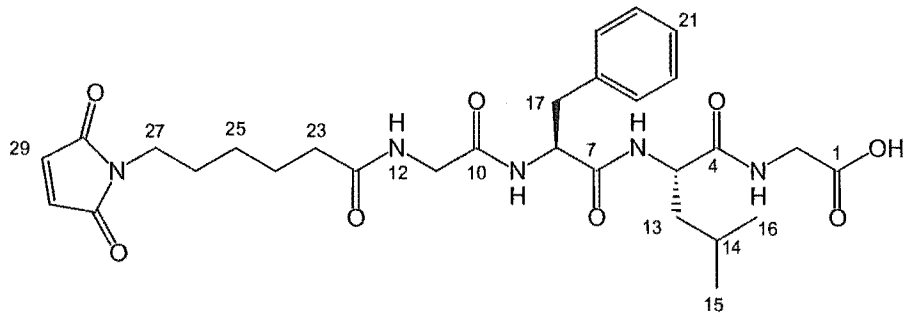
dissolved in MeOH and purified by preparative HPLC (Shimadzu VP system, Phenomenex C18 column, 45% CH_3CN/H_2O (0.05% TFA)) with a retention time of 9.4 min (**68**, 1.04 mg, 38%). HRESIMS gave (MH^+) 743.3424 ($C_{39}H_{47}N_6O_9$ required 743.3405).

5.3.2 Thiolated-CV-N with 68

To the thiolated-CV-N formed in section **3.4.2.2** was added **68** (0.27 mg, $3.63x10^{-4}$ mmole) in DMF (20 μL). After 2 hours the solution was concentrated under a flow of N_2 and analysed by LCMS (Zorbax C3). The products eluted at 11.8 min representing thiolated-CV-N with addition of a single coumarin adduct (11,853 Da) and addition of two coumarin adducts (12,696 Da).

5.4.1 Synthesis of Maleimido-Caproic-Gly-Phe-Leu-Gly-OH

The method for SPPS of **69** was identical to that of **49** (section **4.3.1.2**), but starting with 0.118 meq. of Fmoc-Gly-resin. This required, in turn, 2 equiv. of each amino acid (0.237 mmole), 2 equiv. of maleimidocaproic acid (**67**, 50 mg, 0.237 mmole) with HBTU (90 mg, 0.237 mmole), and 4 equiv. of DIPEA (93 μL, 0.473 mmole) per coupling. Cleavage conditions and workup of the product was also identical to that of **69**. The correct mass was noted by ESIMS analysis although analytical HPLC (Shimadzu VP system, Phenomenex C18 column, 60% $MeOH/H_2O$ (0.05% TFA)) indicated the presence of several minor impurities. Preparative HPLC using the same solvent mixture (Shimadzu LC-4A, Dynamax C18, 205 nm) purified the product (retention time 18 min, 8.2 mg). ESIMS (-ve mode, $(MH)^-$) gave 584.24 Da ($C_{29}H_{38}N_5O_8$ required 584.28). ^1H NMR was recorded on the Unity 300.



maleimido-caproic-Gly-Phe-Leu-Gly-OH (**69**)

^{1}H NMR (CD_3OD) δ 0.92 (m, $(H15)_3$ or $(H16)_3$), δ 0.96 (m, $(H16)_3$ or $(H15)_3$), δ 1.13 (m, $(H25)_2$), δ 1.34-1.55 (m, H14), δ 1.34-1.55 (m, $(H13)_2$), δ 1.34-1.55 (m, $(H24)_2$), δ 1.34-1.55 (m, $(H26)_2$), δ 2.03 (m, $(H23)_2$), δ 2.79 (m, H17a), δ 2.99 (m, H17b), δ 3.14 (m, $(H27)_2$), δ 3.50-3.78 (m, $(H11)_2$), δ 3.50-3.78 (m, $(H2)_2$), δ 4.26 (m, H5), δ 4.45 (m, H8), δ 6.61 (s, 2x H29), δ 7.07 (m, 2x H19), δ 7.07 (m, 2x H20), δ 7.07 (m, H21), amide protons were not observed.

5.4.2 Synthesis of Maleimido-Caproic-Gly-Phe-Leu-Gly-Toxin

To **69** (1.07 mg, 1.83×10^{-3} mmole) in dry THF (1 mL, 3 drops of dry DMF for dissolution) was added DCC (0.75 mg, 3.66×10^{-3} mmole) and NHS (0.42 mg, 3.66×10^{-3} mmole) and the reaction left under argon. ESIMS analysis after 24 hours showed approximately 50% formation of the succinimidyl ester at which stage norhomohalichondrin B amine (**28**, 1 mg, 9.16×10^{-4} mmole) in CH_3CN (100 μL) was added. ESIMS analysis after a total reaction time of 60 hours revealed complete disappearance of the toxin amine and formation of the desired product. Preparative HPLC (Shimadzu VP system, Phenomenex C18 column, 25%-85% CH_3CN/H_2O over 20 min, 210 nm) purified the product (retention time 19.2 min) which was collected and dried. (**71**, 0.84 mg, 55%). HRESIMS gave $(MNa_2)^{2+}$ 852.4125 ($C_{89}H_{122}N_6O_{24}Na_2/2$) required 852.4153).

5.5.1 Maleimide-Toxin Reaction with CV-N-Cys

To CV-N-Cys (**72**, from LDDRD, 0.100 mg, $9.00x10^{-6}$ mmole) in PBS buffer (274 μL) was added **71** (0.150 mg, $9.04x10^{-5}$ mmole) in DMF (74 μL). LCMS analysis (Zorbax C3) after 24 hours indicated formation of approximately 15% product so the reaction was left for a further 24 hours at which time no perceivable increase in the yield was observed. Preparative HPLC (Shimadzu VP system, Zorbax C3 column, 20%-60% CH_3CN/H_2O (0.001% formic) over 20 min) separated the CV-N-Cys (retention time 9.5 min) from the CV-N-Cys-biolinker-toxin construct (retention time 12.4 min). LCMS data shown in **Figure 5.8**.

5.5.2 Maleimide-Toxin Reaction with Thiolated-CV-N

To the thiolated-CV-N (0.100 mg, $9.08x10^{-6}$ mmole, in water) formed in section **3.4.2.3** was added **71** (0.160 mg, $9.64x10^{-5}$ mmole) in DMF (25 μL). LCMS analysis after 90 min indicated the presence of both thiolated-CV-N and the maleimide-toxin construct starting materials along with a small amount of product. The reaction was left for a further 24 hours. LCMS analysis at this stage revealed no maleimide-toxin construct present so purification was by HPLC (Shimadzu VP system, Zorbax C3 column, 20%-60% CH_3CN/H_2O (0.001% formic) over 20 min). The CV-N-toxin eluted with a retention time of 12.9 min. LCMS data shown in **Figure 5.9**.

8.6 Work Described in Chapter 6

Work in this chapter from sections' **6.3.1**, **6.3.2** and **6.3.3** was carried out while a guest at the LDDRD, NCI, Frederick, Maryland, USA.

ELISA and XTT assays were carried out as described in section **8.1** with the following modifications and changes described below.

6.3.1 Initial ELISA Testing

This ELISA was performed without the presence of gp120 in the wells. The CV-N (**15**) or CV-N-conjugates were bound directly to the plate and the rest of the assay was accomplished as per the standard procedure. The response from the CV-N-conjugate **36** was comparable to that of native CV-N (**15**) while **42** was much less effective at binding to gp120 due to lack of protein material in the sample.

Before addition of the substrate buffer the plates were read at 405 nm by the plate reader to ascertain if any contribution from the dye would interfere with the substrate-mediated colour change. The response from the samples was comparable to the background level of the BSA loaded cells therefore confirming no interference. As a side issue the plates were tested for an increasing fluorescence response as the concentration increased which was duly observed.

6.3.2 ELISA Testing of 5-SFX and BODIPY CV-N-Conjugates

The following data (**Table 8.1**) shows the volume and concentration of the CV-N-dye conjugates tested in the ELISA assay and the volume of sample used for the ELISA assay. The volumes for the ELISA assay were made up to 1,000 μL with PBS to give standard

solutions at 5 ng/mL. These were subsequently diluted out for each well to give the amounts shown in **Figure 6.2**.

Table 8.1 CV-N-dye conjugates for ELISA assay

Samples	Volume (mL)	Concentration (μg/mL)	Volume for ELISA (μL)
33	7.7	26.0	192
34	5.0	40.0	125
35	5.9	33.9	147
36	6.2	32.3	154
43	5.0	40.0	125
44	7.0	28.6	175
41	5.8	34.5	145
42	7.0	28.6	175

6.3.3 XTT Assay Testing of CV-N-Dye-Conjugates

The same samples used above for the ELISA assay were also submitted for XTT assay analysis. 200 μL aliquots of each were submitted at their respective concentrations and the EC_{50} values obtained and scaled as below (**Table 8.2**).

Table 8.2 CV-N-dye conjugate EC_{50} values from XTT assay

Samples	Original EC_{50} (μg/mL)	High Test Conc. (μg/mL)	Corrected EC_{50} (μg/mL)
33	0.0015	1.73	0.0026
34	0.0032	2.67	0.0085
35	0.0024	2.26	0.0054
36	0.0016	2.15	0.0034
43	0.0024	2.67	0.0064
44	0.0058	1.90	0.0110
41	0.0390	2.30	0.0897
42	0.6370	1.90	1.2133

6.4.1 XTT Assay Testing of CV-N-Coumarin Conjugates

Samples for this XTT assay were submitted at the same concentration as native CV-N (**15**, 1.1 μg/mL) providing EC_{50} values in **Table 6.2**.

References

1. (UNAIDS), Joint United Nations Programme on HIV/AIDS. and (WHO), World Health Organisation, "*AIDS epidemic update: December 2000.*" **2000**, Geneva. p. 3-10.
2. Vallee, H. and Carre, H., "*Sur l'anemie infectieuse du cheval.*" C. R. Acad. Sci., **1904**, *139*, 1239-1241.
3. Haase, A. T., "*Pathogenesis of lentivirus infections.*" Nature, **1986**, *322*, 130-136.
4. Levy, J. A., "*The multifaceted retrovirus.*" Cancer Res., **1986**, *46*, 5457-5468.
5. Pepin, J.; Morgan, G.; Dunn, D.; Gevao, S.; Mendy, M.; Gaye, I.; Scollen, N.; Tedder, R. and Whittle, H., "*HIV-2-induced immunosuppression among asymptomatic west African prostitutes: Evidence that HIV-2 is pathogenic, but less so than HIV-1.*" AIDS, **1991**, *5*, 1165-1172.
6. Markowitz, D. M., "*Infection with the human immunodeficiency virus type 2.*" Ann. Intern. Med., **1993**, *118*, 211-218.
7. Gottlieb, M. S.; Schroff, R.; Schanker, H. M.; Weisman, J. D.; Fan, P. T.; Wolf, R. A. and Saxon, A., "*Pneumocystis carinii pneumonia and mucosal candidiasis in previously healthy homosexual men: evidence of a new acquired cellular immunodeficiency.*" New Engl. J. Med., **1981**, *305*, 1425-431.
8. Masur, H.; Michelis, M. A.; Greene, J. V.; Onorato, F.; Vande Stouwe, R. A.; Holzman, R. S.; Wormser, G.; Brettman, L.; Lange, M.; Murray, H. W. and Cunningham-Rundles, S., "*An outbreak of community-acquired Pneumocystis carinii pneumonia: initial manifestation of cellular immune dysfunction.*" New Engl. J. Med., **1981**, *305*, 1431-1438.
9. Siegal, F. P.; Lopez, C.; Hammer, G. S.; Brown, A. E.; Kornfeld, S. J.; Gold, J.; Hassett, J.; Hirschman, S. Z.; Cunningham-Rundles, C. and Adelsberg, B. R., "*Severe acquired immunodeficiency in male homosexuals, manifested by chronic perianal ulcerative herpes simplex lesions.*" New Engl. J. Med., **1981**, *305*, 1439-1444.

10. Notkins, A.; Mergenhagen, S. and Howard, R., "*Effect of virus infections on the function of the immune system.*" Annu. Rev. Microbiol., **1970**, *24*, 525-537.

11. Rouse, B. T. and Horohov, D. W., "*Immunosuppression in viral infections.*" Rev. Infect. Dis., **1986**, *8*, 850-873.

12. Chiu, I. M.; Yaniv, A.; Dahlberg, J. E.; Gazit, A.; Skuntz, S. F.; Tronick, S. R. and Aaronson, S. A., "*Nucleotide sequence evidence for relationship of AIDS retrovirus to lentiviruses.*" Nature, **1985**, *317*, 366-368.

13. Gonda, M.; Wong-Staal, F.; Gallo, R. C.; Clements, J. E.; Narayan, O. and Gilden, R. V., "*Sequence homology and morphologic similarity of HTLV-III and visna virus, a pathogenic lentivirus.*" Science, **1985**, *227*, 173-177.

14. Levy, J. A.; Kaminsky, L. S.; Morrow, W. J. W.; Steimer, K.; Luciw, P.; Dina, D.; Hoxie, J. and Oshiro, L., "*Infection by the retrovirus associated with the acquired immunodeficiency syndrome.*" Ann. Intern. Med., **1985**, *103*, 694-699.

15. Rabson, A. and Martin, M., "*Molecular organisation of the AIDS retrovirus.*" Cell, **1985**, *40*, 477-480.

16. Wyatt, R. and Sodroski, J. G., "*The HIV-1 envelope glycoproteins : fusogens, antigens, and immunogens.*" Science, **1998**, *280*, 1884-1888.

17. Gelderblom, H.; Ozel, M.; Hausmann, E. H. S.; Winkel, T.; Pauli, G. and Koch, M. A., "*Fine structure of human immunodeficiency virus (HIV), immunolocalization of structural proteins and virus-cell relation.*" Micron Microscop., **1988**, *19*, 41-60.

18. Gelderblom, H.; Ozel, M. and Pauli, G., "*Morphogenesis and morphology of HIV. Structure-function relations.*" Arch. Virol., **1989**, *106*, 1-13.

19. Layne, S. P.; Merges, M. J.; Dembo, M.; Spouge, J. L.; Conley, S. R.; Moore, J. P.; Raina, J. L.; Renz, H.; Gelderblom, H. and Nara, P. L., "*Factors underlying spontaneous inactivation and susceptibility to neutralization of human immunodeficiency virus.*" Virology, **1992**, *189*, 695-714.

20. Hahn, B. H.; Shaw, G. M.; De Cock, K. M. and Sharp, P. M., "*AIDS as a zoonosis: scientific and health implications.*" Science, **2000**, *287*, 607-614.

21. Gao, F.; Bailes, E. and Robertson, D. L., "*Origin of HIV-1 in the chimpanzee, Pan troglodytes troglodytes.*" Nature, **1999**, *397*, 436-441.

22. De Cock, K. M. and Weiss, H. A., "*The global epidemiology of AIDS.*" Trop. Med. Intl. Health, **2000**, *5*, A3-A9.

23. Peeters, M.; Gueye, A.; Mboup, S.; Bibollet-Ruche, F.; Ekaza, E.; Mulanga, C.; Ouedrago, R.; Gandji, R.; Mpele, P.; Dibanga, G.; Koumare, B.; Saidou, M.; Esu-Williams, E.; Lombart, J. P.; Badombena, W.; Luo, N.; Vanden Haesevelde, M. and Delaporte, E., "*Geographical distribution of HIV-1 group O viruses in Africa.*" AIDS, **1997**, *11*, 493-498.

24. Weiss, R. A. and Wrangham, R. W., "*The origin HIV-1: From Pan to pandemic.*" Nature, **1999**, *397*, 385-386.

25. Weiss, R. A., "*Is AIDS man-made?*" Science, **1999**, *286*, 1305-1306.

26. Saag, M. S.; Holodniy, M.; Kuritzkes, D. R.; O'Brien, W. A.; Coombs, R.; Poscher, M. E.; Jacobsen, D. M.; Shaw, G. M.; Richman, D. D. and Volberding, P. A., "*HIV viral load markers in clinical practice.*" Nature Med., **1996**, *2*, 625-629.

27. Aranout, R. A.; Lloyd, A. L.; O'Brien, T. R.; Goedert, J. J.; Leonard, J. M. and Nowak, M. A., "*A simple relationship between viral load and survival time in HIV-1 infection.*" Proc. Natl. Acad. Sci. U.S.A., **1999**, *96*, 11549-11553.

28. Mellors, J. W.; Rinaldo, C. R.; Gupta, P.; White, R. M.; Todd, J. A. and Kingsley, L. A., "*Prognosis in HIV-1 infection predicted by the quantity of virus in plasma.*" Science, **1996**, *272*, 1167-1170.

29. Perelson, A. S.; Neumann, A. U.; Markowitz, M.; Leonard, J. M. and Ho, D. D., "*HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time.*" Science, **1996**, *271*, 1582-1586.

30. Ho, D. D.; Neumann, A. U.; Perelson, A. S.; Chen, W.; Leonard, J. M. and Markowitz, M., "*Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection.*" Nature, **1995**, *373*, 123-126.

31. Wei, X.; Ghosh, S. K.; Taylor, M. E.; Johnson, V. A.; Emini, E. A.; Deutsch, P.; Lifson, J. D.; Bonhoeffer, S.; Nowak, M. A.; Hahn, B. H.; Saag, M. S. and Shaw, G. M., "*Viral dynamics in human immunodeficiency virus type 1 infection.*" Nature, **1995**, *373*, 117-122.

32. Levy, J. A., "*HIV and the pathogenesis of AIDS.*" 2 ed. **1998**, Washington D. C.: American Society for Microbiology.

33. Haase, A. T.; Henry, K.; Zupancic, M.; Sedgewick, G.; Faust, R. A.; Melroe, H.; Cavert, W.; Gebhard, K.; Staskus, K.; Zhang, Z. Q.; Dailey, P. J.; Balfour, H. H.; Erice, A. and Perelson, A. S., "*Quantitative image analysis of HIV-1 infection in lymphoid tissue.*" Science, **1996**, *274*, 985-989.

34. Maddon, P. J.; Dalgleish, A. G.; McDougal, J. S.; Clapham, P. R.; Weiss, R. A. and Axel, R., "*The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain.*" Cell, **1986**, *11*, 466-474.

35. Feng, Y.; Broder, C. C.; Kennedy, P. E. and Berger, P. E., "*HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor.*" Science, **1996**, *272*, 872-877.

36. Deng, H. K.; Liu, R.; Ellmeier, W.; Choe, S.; Unutmaz, D.; Burkhart, M.; Marzio, P. D.; Marmon, S.; Sutton, R. E.; Hill, C. M.; Davis, C. B.; Peiper, S. C.; Schall, T. J.; Littman, D. R. and Landau, N. R., "*Identification of a major co-receptor for primary isolates of HIV-1.*" Nature, **1996**, *381*, 661-666.

37. Dragic, T.; Litwin, V.; Allaway, G. P.; Martin, S. R.; Huang, Y.; Nagashima, K. A.; Cayanan, C.; Maddon, P. J.; Koup, R. A.; Moore, J. P. and Paxton, W. A., "*HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5.*" Nature, **1996**, *381*, 667-673.

38. Berger, E. A.; Doms, R. W.; Fenyo, E. M.; Korber, B. T. M.; Littman, D. R.; Moore, J. P.; Sattentau, Q. J.; Schuitemaker, H.; Sodroski, J. and Weiss, R. A., "*A new classification for HIV-1.*" Nature, **1998**, *391*, 240.

39. Michael, N. L., "*Host genetic influences on HIV-1 pathogenesis.*" Curr. Opin. in Immun., **1999**, *11*, 466-474.

40. Carrington, M.; Dean, M.; Martin, M. P. and O'Brien, S. J., "*Genetics of HIV-1 infection: chemokine receptor CCR5 polymorphism and its consequences.*" Hum. Mol. Genetics, **1999**, *8*, 1939-1945.

41. Samson, M.; Libert, F.; Doranz, B. J.; Rucker, J.; Liesnard, C.; Farber, C. M.; Saragosti, S.; Lapoumeroulie, C.; Cognaux, J.; Forceille, C.; Muyldermans, G.; Verhofstede, C.; Burtonboy, G.; Georges, M.; Imai, T.; Rana, S.; Yi, Y.; Smyth, R. J.; Collman, R. G.; Doms, R. W.; Vassart, G. and Parmentier, M., "*Resistance to*

HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene." Nature, **1996**, *382*, 722-725.

42. Weiss, R. A., "*Getting to know HIV.*" Trop. Med. Intl. Health, **2000**, *5*(7), A10-A15.
43. Ugolini, S., "*HIV-1 attachment: another look.*" Trends Microbiol., **1999**, *7*, 144-149.
44. Kwong, P. D.; Wyatt, R.; Robinson, J.; Sweet, R. W.; Sodroski, J. G. and Hendrickson, W. A., "*Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody.*" Nature, **1998**, *393*, 648-659.
45. Baba, M.; Nishimura, O.; Kanzaki, N.; Okamoto, M.; Sawada, H.; Iizawa, Y.; Shiraishi, M.; Aramaki, Y.; Okonogi, K.; Ogawa, Y.; Meguro, K. and Fujino, M., "*A small-molecule, nonpeptide CCR5 antagonist with highly potent and selective anti-HIV-1 activity.*" Proc. Natl. Acad. Sci. U.S.A., **1999**, *96*, 5698-5703.
46. Donzella, G. A.; Schols, D.; Lin, S. W.; Este, J. A.; Nagashima, K. A.; Maddon, P. J.; Allaway, G. P.; Sakamar, T. P.; Henson, G.; De Clereq, E. and Moore, J. P., "*AMD3100, a small molecule inhibitor of HIV-1 entry via the CXCR4 co-receptor.*" Nature Med., **1998**, *4*, 72-77.
47. Doranz, B. J.; Grovit-Ferbas, K.; Sharron, M. P.; Mao, S. H.; Goetz, M. B.; Daar, E. S.; Doms, R. W. and O'Brien, W. A., "*A small-molecule inhibitor directed against the chemokine receptor CXCR4 prevents its use as an HIV-1 co-receptor.*" J. Exp. Med., **1997**, *186*, 1395-1400.
48. Murakami, T.; Nakajima, T.; Koyanagi, Y.; Tachibana, K.; Fujii, N.; Tamamura, H.; Yoshida, N.; Waki, M.; Matsumoto, A.; Yoshie, O.; Kishimoto, T.; Yamamoto, N. and Nagasawa, T., "*A small molecule CXCR4 inhibitor that blocks T cell line-tropic HIV-1 infection.*" J. Exp. Med., **1997**, *186*, 1389-93.
49. Vartanian, J. P., "*AMD-3100. AnorMED.*" IDrugs, **2000**, *3*(7), 811-816.
50. Pommier, Y.; Marchand, C. and Neamati, N., "*Retroviral integrase inhibitors year 2000: update and perspectives.*" Antiviral Res., **2000**, *47*(3), 139-148.
51. Turpin, J. A.; Song, Y.; Inman, J. K.; Huang, M.; Wallqvist, A.; Maynard, A.; Covell, D. G.; Rice, W. G. and Appella, E., "*Synthesis and biological properties of*

novel pyridinioalkanoyl thiolesters (PATE) as anti-HIV-1 agents that target the viral nucleocapsid protein zinc fingers." J. Med. Chem., **1999**, *42*, 67-86.

52. Druillennec, S.; Dong, C. Z.; Escaich, S.; Gresh, N.; Bousseau, A.; Roques, B. P. and Fournie-Zaluski, M. C., "*A mimic of HIV-1 nucleocapsid protein impairs reverse transcription and displays antiviral activity.*" Proc. Natl. Acad. Sci. U.S.A., **1999**, *96*, 4886-4891.

53. Ezzell, C., "*Emergence of the protease inhibitors: a better class of AIDS drugs?*" J. NIH Res., **1996**, *8*, 41-45.

54. Lam, P. Y.; Jadhav, P. K.; Eyermann, C. J.; Hodge, C. N.; Ru, Y.; Bacheler, L. T.; Meek, J. L.; Otto, M. J.; Rayner, M. M.; Wong, Y. N.; Chang, C. H.; Weber, P. C.; Jackson, D. A.; Sharpe, T. R. and Erickson-Viitanen, S., "*Rational design of potent, bioavailable, nonpeptide cyclic ureas as HIV protease inhibitors.*" Science, **1994**, *263*, 380-384.

55. Autran, B.; Carcelain, G.; Li, T. S.; Blanc, C.; Mathez, D.; Tubiana, R.; Katlama, C.; Debre, P. and Leibowitch, J., "*Positive effects of combined antiretroviral therapy on CD4+ T cell homeostasis and function in advanced HIV disease.*" Science, **1997**, *277*, 112-116.

56. Palella, F. J.; Delaney, K. M.; Moorman, A. C.; Loveless, M. O.; Fuhrer, J.; Satten, G. A.; Aschman, D. J. and Holmberg, S. D., "*Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection.*" New Engl. J. Med., **1998**, *338*, 853-860.

57. Vittinghoff, E.; Scheer, S.; O'Malley, P.; Colfax, G.; Holmberg, S. D. and Buchbinder, S. P., "*Combination antiretroviral therapy and recent declines in AIDS incidence and mortality.*" J. Infect. Dis., **1999**, *179*, 717-720.

58. Dewhurst, S.; da Cruz, R. L. W. and Whetter, L., "*Pathogenesis and treatment of HIV-1 infection: recent developments (Y2K update).*" Front. Biosci., **2000**, *5*, D30-D49.

59. Birch, C., "*Implications of poor compliance.*" J. HIV Ther., **1998**, *3*, 63-66.

60. Siliciano, R. F., "*Latency and reservoirs for HIV-1.*" AIDS, **1999**, *13*, S49-S58.

61. Chun, T. W.; Engel, D.; Mizell, S. B.; Ehler, L. A. and Fauci, A. S., "*Induction of HIV-1 replication in latently infected CD4+ T cells using a combination of cytokines.*" J. Exp. Med., **1998**, *188*, 83-91.

62. Chun, T. W.; Engel, D.; Mizell, S. B.; Hallahan, C. W.; Fischette, M.; Park, S.; Davey, R. T. J.; Dybul, M.; Kovacs, J. A.; Metcalf, J. A.; Mican, J. M.; Berrey, M. M.; Corey, L.; Lane, H. C. and Fauci, A. S., "*Effect of interleukin-2 on the pool of latently infected, resting CD4+ T cells in HIV-1 infected patients receiving highly active anti-retroviral therapy.*" Nature Med., **1999**, *5*, 651-655.

63. Smith, C.; Lilly, S. and Miralles, G. D., "*Treatment of HIV infection with cytoreductive agents.*" AIDS Res. Hum. Retroviruses, **1998**, *14*, 1305-1313.

64. Piot, P., "*The science of AIDS: A tale of two worlds.*" Science, **1998**, *280*, 1844-1845.

65. Francis, D. P.; Gregory, T.; McElrath, M. J.; Belshe, R. B.; Gorse, G. J.; Migasena, S.; Kitayaporn, D.; Pitisuttitham, P.; Matthews, T.; Schwartz, D. H. and Berman, P. W., "*Advancing AIDSVAX to phase 3. Safety, immunogenicity, and plans for phase 3.*" AIDS Res. Hum. Retroviruses, **1998**, *14*, S325-331.

66. Migasena, S.; Suntharasamai, P.; Pitisuttithum, P.; Kitayaporn, D.; Wasi, C.; Huang, W.; Vanichseni, S.; Koompong, C.; Kaewkungwal, J.; Raktham, S.; Ippolito, T.; Hanson, C.; Gregory, T.; Heyward, W. L.; Berman, P. and Francis, D., "*AIDSVAX (MN) in bangkok injecting drug users: a report on safety and immunogenicity, including macrophage-tropic virus neutralization.*" AIDS Res. Hum. Retrovir., **2000**, *16*(7), 655-663.

67. MacGregor, R. R.; Boyer, J. D.; Ugen, K. E.; Lacy, K. E.; Gluckman, S. J.; Bagarazzi, M. L.; Chattergoon, M. A.; Baine, Y.; Higgins, T. J.; Ciccarelli, R. B.; Coney, L. R.; Ginsberg, R. S. and Weiner, D. B., "*First human trial of a DNA-based vaccine for treatment of human immunodeficiency virus type 1 infection: safety and host response.*" J. Infect. Dis., **1998**, *178*(1), 92-100.

68. Fang, Z. Y.; Kuli-Zade, I. and Spearman, P., "*Efficient human immunodeficiency virus (HIV-1) Gag-Env pseudovirion formation elicited from mammalian cells by a canrypox HIV vaccine candidate.*" J. Infect. Dis., **1999**, *180*, 1122-1132.

69. Clements, M. L.; Weinhold, K.; Siliciano, R.; Schwartz, D.; Matthews, T.; Graham, B.; Keefer, M.; McElrath, J.; Gorse, G.; Hsieh, R.; Duliege, A.; Excler, J.; Meigner, B.; Tartaglia, J. and Paoletti, E. "*HIV immunity induced by canarypox (ALVAC)-MN gp160,-SF2 rgp120 or both.*" in *12th World AIDS Conference.* **1998**. Geneva, Switzerland.

70. Patterson, G. M. L.; Baker, K. K.; Baldwin, C. L.; Bolis, C. M.; Caplan, F. R.; Larsen, L. K.; Levine, I. A.; Moore, R. E.; Nelson, C. S.; Tschappat, K. D.; Tuang, G. D.; Boyd, M. R.; Cardellina, J. H. I.; Collins, R. P.; Gustafson, K. R.; Snader, K. M.; Weislow, O. S. and Lewin, R. A., "*Antiviral activity of cultured blue-green algae (cyanophyta).*" J. Phycol., **1993**, *29*, 125-130.

71. Boyd, M. R.; Gustafson, K. R.; McMahon, J. B.; Shoemaker, R. H.; O'Keefe, B. R.; Mori, T.; Gulakowski, R. J.; Wu, L.; Rivera, M. I.; Laurencot, C. M.; Cardellina, J. H. I.; Buckheit, R. W. J.; Nara, P. L.; Pannell, L. K.; Sowder, R. C. I. and Henderson, L. E., "*Discovery of Cyanovirin-N, a novel human immunodeficiency virus-inactivating protein that binds viral surface envelope glycoprotein gp120: potential application to microbicide development.*" Antimicrob. Agents Chemother., **1997**, *41*, 1521-1530.

72. Gustafson, K. R.; Sowder, R. C. I.; Henderson, L. E.; Cardellina, J. H. I.; McMahon, J. B.; Rajamani, U.; Pannell, L. K. and Boyd, M. R., "*Isolation, primary sequence determination, and disulfide bond structure of cyanovirin-N, an anti-HIV (human immunodeficiency virus) protein from the cyanobacterium Nostoc ellipsosporum.*" Biochem. Biophys. Res. Commun., **1997**, *238*, 223-228.

73. Mariner, J. M.; McMahon, J. B.; O'Keefe, B. R.; Nagashima, K. and Boyd, M. R., "*The HIV-inactivating protein, cyanovirin-N, does not block gp120-mediated virus-to-cell binding.*" Biochem. Biophys. Res. Commun., **1998**, *248*, 841-845.

74. Esser, M. T.; Mori, T.; Mondor, I.; Sattentau, Q. J.; Dey, B.; Berger, E. A.; Boyd, M. R. and Lifson, J. D., "*Cyanovirin-N binds to gp120 to interfere with CD4-dependent human immundeficienct type 1 virion binding, fusion, and infectivity but does not effect the CD4 binding site on gp120 or soluble CD4-induced conformational changes in gp120.*" J. Virol., **1999**, *73*, 4360-4371.

75. Trkola, A.; Purtscher, M.; Muster, T.; Ballaun, C.; Buchacher, A.; Sullivan, N.; Srinivasan, K.; Sodroski, J. G.; Moore, J. P. and Katinger, H., "*Human monoclonal antibody 2G12 defines a distinctive neutralization epitope on the gp120 glycoprotein of human immunodeficiency virus type 1.*" J. Virol., **1996**, *70*, 1100-1108.

76. Wyatt, R.; Kwong, P. D.; Desjardins, E.; Sweet, R. W.; Robinson, J.; Hendrickson, W. A. and Sodroski, J. G., "*The antigenic structure of the HIV gp120 envelope glycoprotein.*" Nature, **1998**, *393*, 705-711.

77. Dey, B.; Lerner, D. L.; Lusso, P.; Boyd, M. R.; Elder, J. H. and Berger, E. A., "*Multiple antiviral activities of cyanovirin-N: Blocking of gp120 interaction with CD4 and coreceptor and inhibition of diverse enveloped viruses.*" J. Virol., **2000**, *74*, 4562-4569.

78. O'Keefe, B. R.; Shenoy, S. R.; Xie, D.; Zhang, W.; Muschik, J. M.; Currens, M. J.; Chaiken, I. and Boyd, M. R., "*Analysis of the interaction between the HIV-inactivating protein cyanovirin-N and soluble forms of the envelope glycoproteins gp120 and gp41.*" Mol. Pharmacol., **2000**, *58*, 982-992.

79. Geyer, H.; Holschbach, C.; Hunsmann, G. and Schneider, J., "*Carbohydrates of human immunodeficiency virus: Structures of oligosaccharides linked to the envelope glycoprotein 120.*" J. Biol. Chem., **1988**, *263*, 11760-11767.

80. Mizuochi, T.; Matthews, T.; Kato, M.; Hamako, J.; Titani, K.; Solomon, J. and Feizi, T., "*Diversity of oligosaccharide structures on the envelope glycoprotein gp120 of human immunodeficiency virus 1 from lymphoblastoid cell line H9.*" J. Biol. Chem., **1990**, *265*, 8519-8524.

81. Yeh, J. C.; Seals, J. R.; Murphy, C. I.; van Halbeek, H. and Cummings, R. D., "*Site-specific N-glycosylation and oligosaccharide structures of recombinant HIV-1 gp120 derived from a baculovirus expression system.*" Biochemistry, **1993**, *32*, 11087-11099.

82. Mori, T. and Boyd, M. R., "*Cyanovirin-N, a potent Human Immunodeficiency Virus-inactivating protein, blocks both CD4-dependent and CD4-independent binding of soluble gp120 (sgp120) to target cells, inhibits sCD4-induced binding of*

sgp120 to cell-associated CXCR4, and dissociates bound sgp120 from target cells." Antimicrob. Agents Chemother., **2001**, *45*, 664-672.

83. Mori, T.; Gustafson, K. R.; Pannell, L. K.; Shoemaker, R. H.; Wu, L.; McMahon, J. B. and Boyd, M. R., "*Recombinant production of cyanovirin-N, a potent human immunodeficiency virus-inactivating protein derived from a cultured cyanobacterium.*" Protein Expr. Purif., **1998**, *12*, 151-158.

84. Mori, T.; Shoemaker, R. H.; Gulakowski, R. J.; Krepps, B. L.; McMahon, J. B.; Gustafson, K. R.; Pannell, L. K. and Boyd, M. R., "*Analysis of sequence requirements for biological activity of cyanovirin-N, a potent HIV(human immunodeficiency virus)-inactivating protein.*" Biochem. Biophys. Res. Commun., **1997**, *238*, 218-222.

85. Bewley, C. A.; Gustafson, K. R.; Boyd, M. R.; Covell, D. G.; Bax, A.; Clore, G. M. and Gronenborn, A. M., "*Solution structure of cyanovirin-N, a potent HIV-inactivating protein.*" Nature Struct. Biol., **1998**, *5*, 571-578.

86. Tjandra, N.; Omichinski, J. G.; Gronenborn, A. M.; Clore, G. M. and Bax, A., "*Use of dipolar ^{1}H-^{15}N and ^{1}H-^{13}C couplings in the structure determination of magnetically oriented macromolecules in solution.*" Nature Struct. Biol., **1997**, *4*, 732-738.

87. Jones, S. and Thornton, J. M., "*Principles of protein-protein interactions.*" Proc. Natl. Acad. Sci. U.S.A., **1996**, *93*, 13-20.

88. Covell, D. G.; Smythers, G. W.; Gronenborn, A. M. and Clore, G. M., "*Analysis of hydrophobicity in the α and β chemokine families and its relevance to dimerization.*" Prot. Sci., **1994**, *3*, 2064-2072.

89. Young, L.; Jernigan, R. L. and Covell, D. G., "*A role for surface hydrophobicity in protein-protein recognition.*" Prot. Sci., **1994**, *3*, 717-729.

90. Villoutreix, B. O.; Hardig, Y.; Wallqvist, A.; Covell, D. G.; Frutos, G. d. and Dahlback, B., "*Structural investigation of C4p-binding protein by molecular modeling: localization of putative binding sites.*" Proteins: Struct., Funct., Genet., **1998**, *31*(4), 391-405.

91. Yang, F.; Bewley, C. A.; Louis, J. M.; Gustafson, K. R.; Boyd, M. R.; Gronenborn, A. M.; Clore, G. M. and Wlodawer, A., "*Crystal structure of cyanovirin-N, a potent*

HIV-inactivatiing protein, shows unexpected domain swapping." J. Mol. Biol., **1999**, *288*, 403-412.

92. Bewley, C. A. and Clore, G. M., "*Determination of the relative orientation of the two halves of the domain-swapped dimer of cyanovirin-N in solution using dipolar couplings and rigid body minimization.*" J. Am. Chem. Soc., **2000**, *122*, 6009-6016.
93. Fauci, A. S., "*NIAID expands research on topical microbicides to prevent STD's in women*", in *NIAID news*. **1995**: Office of Communications, NIAID, NIH.
94. Lange, J. M. A.; Karam, M. and Piot, P., "*Boost for vaginal microbicides against HIV.*" The Lancet, **1993**, *341*(1356).
95. Painter, K., "*Researching urgent need for women's and self protection*", in *USA Today*. **1996**. p. 4D.
96. The International Working Group on Vaginal Microbicides., "*Recommendations for the development of vaginal microbicides.*" AIDS, **1996**, *10*, 1-6.
97. Dean, M.; Carrington, M.; Winkler, C.; Huttley, G. A.; Smith, M. W.; Allikmets, R.; Goedert, J. J.; Buchbinder, S. P.; Vittinghoff, E.; Gomperts, E.; Donfield, S.; Vlahov, D.; Kaslow, R.; Saah, A.; Rinaldo, C. R. and Detels, R., "*Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene.*" Science, **1996**, *273*, 1856-1862.
98. Hill, C. M. and Littman, D. R., "*Natural reistance to HIV?*" Nature, **1996**, *382*, 688-699.
99. Liu, R.; Paxton, W. A.; Choe, S.; Ceradini, D.; Martin, S. R.; Horuk, R.; MacDonald, M. E.; Stuhlmann, H.; Koup, R. A. and Landau, N. R., "*Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection.*" Cell, **1996**, *86*, 367-377.
100. Daar, E. S.; Li, X. L.; Moudgli, T. and Ho, D., "*High concentrations of recombinant soluble CD4 are required to neutralize primary human immunodeficiency virus type 1 isolates.*" Proc. Natl. Acad. Sci. U.S.A., **1990**, *87*, 6574-6578.
101. Moore, J. P.; Mckeating, J. A.; Huang, Y.; Ashkenazi, A. and Ho, D. O., "*Virions of primary human immunodeficiency virus type 1 isolates resistant to soluble CD4*

(sCD4) neutralization differ in sCD4 binding and glycoprotein gp120 retention from sCD4-sensitive isolates." J. Virol., **1992**, *66*, 235-243.

102. Orloff, S. L.; Bandea, C. L.; Kennedy, M. S.; Allaway, G. P.; Maddon, P. J. and McDougal, J. S., "*Increase in sensitivity to soluble CD4 by primary HIV type 1 isolates after passage through C8166 cells: Association with sequence differences in the first constant (C1) region of glycoprotein 120.*" AIDS Res. Hum. Retrovir., **1995**, *11*, 335-342.

103. Schooley, R. T.; Merigan, T. C.; Gaut, P.; Hirsch, M. S.; Holodniy, M.; Flynn, T.; Liu, S.; Byington, R. E.; Henochowicz, S.; Gubish, E.; Spriggs, D.; Kufe, D.; Schindler, J.; Dawson, A.; Thomas, D.; Hanson, D. G.; Letwin, B.; Liu, T.; Gulinello, J.; Kennedy, S.; Fisher, R. and Ho, D. D., "*Recombinant soluble CD4 therapy in patients with the acquired immunodeficiency syndrome (AIDS) and AIDS-related complex. A phase I-II escalating dosage trial.*" Ann. Intern. Med., **1990**, *112*, 247-253.

104. Husson, R. N.; Chung, Y.; Mordenti, J.; Butler, K. M.; Chen, S.; Duiege, A. M.; Brouwers, P.; Jarosinski, P.; Mueller, B. U.; Ammann, A. and Pizzo, P. A., "*Phase I study of continuous-infusion soluble CD4 as a single agent and in combination with oral dideoxyinosine therapy in children with symptomatic human immunodeficiency virus infection.*" J. Pediatr., **1992**, *121*, 627-633.

105. Langner, K. D.; Niedrig, M.; Fultz, P.; Anderson, D.; Reiner, G.; Repke, H.; Gelderblom, H.; Seed, B.; Hilfenhaus, J. and Zelltmeissl, G., "*Antiviral effects of different CD4-immunoglobulin constructs against HIV-1 and SIV: Immunological characterization, pharmacokinetic data and in vivo experiments.*" Arch. Virol., **1993**, *130*(157-170).

106. Davey, R. T. J.; Boenning, C. M.; Herpin, B. R.; Batts, D. H.; Metcalf, J. A.; Wathen, L.; Cox, S. R.; Polis, M. A.; Kovacs, J. A.; Falloon, J.; Walker, R. E.; Salzman, N.; Masur, H. and Lane, H. C., "*Use of recombinant soluble CD4 Pseudomonas exotoxin, a novel immunotoxin, for treatment of persons infected with human immunodeficiency virus.*" J. Infect. Dis., **1994**, *170*, 1180-1188.

107. Ramachandran, R. V.; Katzenstein, D. A.; Wood, R.; Batts, D. H. and Merigan, T. C., "*Failure of short-term CD4-PE40 infusions to reduce virus load in human immunodeficiency virus-infected persons.*" J. Infect. Dis., **1994**, *170*, 1009-1013.

108. Mori, T.; Shoemaker, R. H.; McMahon, J. B.; Gulakowski, R. J.; Gustafson, K. R. and Boyd, M. R., "*Construction and enhanced cytotoxicity of a [cyanovirin-N]-[Pseudomonas exotoxin] conjugate against human immunodeficiency virus-infected cells.*" Biochem. Biophys. Res. Commun., **1997**, *239*, 884-888.

109. Hwang, J.; Fitzgerald, D. J. P.; Adhya, S. and Pastan, I., "*Functional domains of Pseudomonas exotoxin identified by deletion analysis of the gene expressed in E. coli.*" Cell, **1987**, *48*, 129-136.

110. Kreitman, R. J.; Siegall, C. B.; Chaudhary, V. K.; Fitzgerald, D. J. P. and Pastan, I., "*Properties of chimeric toxins with two recognition domains: interleukin 6 and transforming growth factor* α *at different locations in Pseudomonas exotoxin.*" Bioconjugate Chem., **1992**, *3*, 63-68.

111. Gandhi, M. J.; Boyd, M. R.; Yi, L.; Yang, G. G. and Vyas, G. N., "*Properties of cyanovirin-N (CV-N): inactivation of HIV-1 by sessile cyanovirin-N (sCV-N).*" Advances in Transfusion Safety. Dev. Biol., **2000**, *102*, 141-148.

112. McMahon, J. B.; Beutler, J. A.; O'Keefe, B. R.; Goodrum, C. B. B.; Myers, M. A. and Boyd, M. R., "*Development of a cynaovirin-N-HIV-1 gp120 binding assay for high throughput screening of natural product extracts by time-resolved fluorescence.*" J. Biomol. Screen., **2000**, *5*, 169-176.

113. De Clereq, E., "*Antiviral agents: Characteristic activity spectrum depending on the molecular target with which they interact.*" Adv. Viral Res., **1993**, *42*, 1-55.

114. Beutler, J. A.; McKee, T. C.; Fuller, R. W.; Tischler, M.; Cardellina, J. H. I.; Snader, K. M.; McCloud, T. G. and Boyd, M. R., "*Frequent occurrence of HIV-inhibitory sulphated polysaccharides in marine invertebrates.*" Antiviral Chem. Chemother., **1993**, *4*, 167-172.

115. Williams, D. H.; Stone, M. J.; Hauck, P. R. and Rahman, S. K., "*Why are secondary metabolites (natural products) biosynthesized?*" J. Nat. Prod., **1989**, *52*, 1189-1208.

116. Sarma, A. S.; Daum, T. and Muller, W. E. G., "*Secondary metabolites from marine sponges.*" **1993**, Berlin: Akademie gemeinntziger Wissenschaften zu Erfut.

117. Perry, N. B.; Blunt, J. W.; Munro, M. H. G. and Thompson, A. M., "*Discorhabdin C, a highly cytotoxic pigment from a sponge of the genus Latrunculia.*" J. Org. Chem, **1986**, *51*(26), 5476-5478.

118. Perry, N. B.; Blunt, J. W.; Munro, M. H. G. and Thompson, A. M., "*Antiviral and antitumor agents from a New Zealand sponge, Mycale sp. 2. Structures and solution conformations of mycalamides A and B.*" J. Org. Chem, **1990**, *55*(1), 223-227.

119. Hirata, Y. and Uemura, D., "*Halichondrins - anitumor polyether macrolides from a marine sponge.*" Pure Appl. Chem., **1986**, *58*, 701-710.

120. Uemura, D.; Takahashi, K.; Yamamoto, T.; Katayama, C.; Tanaka, J.; Okumura, Y. and Hirata, Y., "*Norhalichondrin A: an antitumor polyether macrolide from a marine sponge.*" J. Am. Chem. Soc., **1985**, *107*, 4796-4798.

121. Pettit, G. R.; Herald, C. L.; Boyd, M. R.; Leet, J. E.; Dufresne, C.; Doubek, D. L.; Schmidt, J. M.; Cerny, R. L.; Hooper, J. N. A. and Rützler, K. C., "*Antineoplastic agents. 219. Isolation and structure of the cell growth inhibitory constituents from the western Pacific marine sponge Axinella sp.*" J. Med. Chem, **1991**, *34*, 3339-3340.

122. Pettit, G. R.; Tan, R.; Williams, M. D.; Doubek, D. L.; Boyd, M. R.; Schmidt, J. M.; Chapius, J. C.; Hamel, E.; Bai, R.; Hooper, J. N. A. and Tackett, L. P., "*Isolation and structure of halistatin 1 from the eastern Indian Ocean marine sponge Phakellia carteri.*" J. Org. Chem, **1993**, *58*, 2538-2543.

123. Pettit, G. R.; Feng, G.; Doubek, D. L.; Boyd, M. R.; Hamel, E.; Bai, R.; Schmidt, J. M.; Tackett, L. P. and Rützler, K. C., "*Antineoplastic agents. CCLII. Isolation and structure of halistatin 2 from the comoros marine sponge Axinella carteri.*" Gazetta Chimica Italiana, **1993**, *123*, 371-377.

124. Lake, R. J., *Internal Report dated 26/2/88.* **1988**, University of Canterbury.

125. Berquist, P. R., "*Sponges.*" 1 ed, Hutchinson & Co. Ltd., Editors. **1978**, London. p14.

126. Paull, K.; Lin, C.; Malspeis, L. and Hamel, E., "*Identification of novel antimitotic agents acting at the tubulin level by computer-assisted evaluation of differential cytotoxicity data.*" Cancer Res., **1992**, *52*, 3892-3900.

127. Bai, R.; Paull, K. D.; Herald, C. L.; Malspeos, L.; Pettit, G. R. and Hamel, E. J., "*Halichondrin B and homohalichondrin B, marine natural products binding in the vinca domain of tubulin. Discovery of tubulin-based mechanism of action by analysis of differential cytotoxicity data.*" J. Biol. Chem., **1991**, *266*, 15882-15889.

128. Munro, M. H. G.; Blunt, J. W.; Lake, R. J.; Litaudon, M.; Battershill, C. N. and Page, M. J., "*Sponges in Time and Space*", R.W.M. van Soest, T.M.G. van Kempen, and J.C. Braekman, Editors. **1994**, Balkema: Rotterdam. p. 473-484.

129. Dumdei, E. J.; Blunt, J. W.; Munro, M. H. G.; Battershill, C. N. and Page, M. J., "*Sponge Sciences. Multidisciplinary Perspectives*", Y. Watanabe and N. Fusetani, Editors. **1998**, Springer-Verlag: Tokyo. p. 353-364.

130. Lill, R. E., "*Studies on New Zealand Marine Natural Products*", **1999**, PhD Thesis, University of Canterbury: Christchurch.

131. Norcross, R. D. and Paterson, I., "*Total Synthesis of Bioactive Marine Macrolides.*" Chem. Rev., **1995**, *95*, 2041-2114.

132. Aicher, T. D.; Buszek, K.; Fang, F. G.; Forsyth, C. J.; Jung, S. H.; Kishi, Y.; Matelich, M. C.; Scola, P. M.; Spero, D. M. and Yoon, S. K., "*Total synthesis of halichondrin B and norhalichondrin B.*" J. Am. Chem. Soc., **1992**, *114*, 3162-3164.

133. Stamos, D. P. and Kishi, Y., "*Synthetic studies on halichondrins: a practical synthesis of the C.1-C.13 segment.*" Tet. Lett., **1996**, *37*, 8643-8646.

134. Stamos, D. P.; Chen, S. S. and Kishi, Y., "*New Synthetic Route to the C.14-C.38 Segment of Halichondrins.*" J. Org. Chem, **1997**, *62*, 7552-7553.

135. Pincus, S. H., "*Targeting HIV-infected cells.*" Methods Mol. Biol., **2000**, *25*, 193-214.

136. Thrush, G. R.; Lark, L. R.; Clinchy, B. C. and Vitetta, E. S., "*Immunotoxins: an update.*" Annu. Rev. Immunol., **1996**, *14*, 49-71.

137. Pincus, S. H., "*Therapeutic potential of anti-HIV immunotoxins.*" Antiviral Res., **1996**, *33*, 1-9.

138. Pastan, I.; Chaudhary, V. K. and Fitzgerald, D. J. P., "*Recombinant toxins as novel therapeutic agents.*" Annu. Rev. Biochem., **1992**, *61*, 331-354.

139. Berger, E. A.; Moss, B. and Pastan, I., "*Reconsidering targeted toxins to eliminate HIV infection: You gotta have HAART.*" Proc. Natl. Acad. Sci. U.S.A., **1998**, *95*, 11511-11513.

140. Gulakowski, R. J.; McMahon, J. B.; Buckheit, R. W. J.; Gustafson, K. R. and Boyd, M. R., "*Antireplicative and anticytopathic activities of prostratin, a non-tumor-promoting phorbol ester, against human immunodeficiency virus (HIV).*" Antiviral Res. Vol. 33. **1997**. 87-97.

141. Duncan, R., "*Drug-polymer conjugates: potential for improved chemotherapy.*" Anti-Cancer Drugs, **1992**, *3*, 175-210.

142. Lundblad, R. and Bradshaw, R., "*Applications of site-specific chemical modification in the manufacture of biopharmaceuticals: I. An overview.*" Biotechnol. Appl. Biochem., **1997**, *26*, 143-151.

143. Takeda, K.; Akagi, Y.; Saiki, A.; Tsukahara, T. and Ogura, H., "*Convenient methods for syntheses of active carbamates, ureas and nitrosoureas using N,N'-disuccinimido carbonate (DSC).*" Tet. Lett., **1983**, *24*, 4569-4572.

144. Ishikawa, K. and Hirata, H., "*Chemistry of succinimido esters XI. Chemical modification of Candida cylindracea lipase with N-succinimidyl carboxylates.*" J. Jpn. Oil Chem. Soc., **1988**, *37*, 458-460.

145. Ishikawa, K. and Hirata, H., "*Chemistry of succinimido esters. XIII. Improvement of Maltosidase activity of Porcine Pancreatic α-Amylase by chemical modification with N-acyloxysuccinimides.*" J. Jpn. Oil Chem. Soc., **1989**, *38*, 60-64.

146. Hirata, H.; Yamashina, T. and Higuchi, K., "*Chemistry of succinimido esters. XIX. N-arylacetylation of amino groups of Bovine Serum Albumin with N-succinimidyl arylacetates.*" J. Jpn. Oil Chem. Soc., **1991**, *40*, 406-414.

147. Morpurgo, M.; Bayer, E. A. and Wilchek, M., "*N-hydroxysuccinimide carbonates and carbamates are useful reactive reagents for coupling ligands to lysines on proteins.*" J. Biochem. Biophys. Methods, **1999**, *38*, 17-28.

148. Shimada, K. and Mitamura, K., "*Derivatization of thiol-containing compounds.*" J. Chromatogr., B: Biomed. Appl., **1994**, *659*, 227-241.

149. Haugland, R. P., "*Handbook of fluorescent probes and research chemicals.*" 6th ed. **1996**, Eugene, Oregon: Molecular Probes.

150. Jue, R.; Lambert, J. M.; Pierce, L. R. and Traut, R. R., "*Addition of sulfhydryl groups of Escherichia coli ribosomes by protein modification with 2-iminothiolane (methyl 4-mercaptobutyrimidate).*" Biochemistry, **1978**, *17*, 5399-5406.

151. Squire, M. A., "*Modification of a marine natural product*", **1999**, Honours project report, University of Canterbury: Christchurch.

152. Devenish, S. R. A., Personal communication, **2000**.

153. Bodansky, M., "*Peptide Chemistry - A Practical Textbook.*" 2nd ed. **1993**, Heidelberg: Springer-Verlag.

154. Fields, G. B. and Noble, R. L., "*Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids.*" Int. J. Peptide Protein Res., **1990**, *35*, 161-214.

155. Alves, L. C.; Almeida, P. C.; Franzoni, L.; Juliano, L. and Juliano, M. A., "*Synthesis of Nα-protected aminoacyl 7-amino-4-methylcoumarin amide by phosphorous oxychloride and preparation of specific fluorogenic substrates for papain.*" Peptide Research, **1996**, *9*, 92-96.

156. Tamamura, H.; Ishihara, T.; Oyake, H.; Imai, M.; Otaka, A.; Ibuka, T.; Arakaki, R.; Nakashima, H.; Murakami, T.; Waki, M.; Matsumoto, A.; Yamamoto, N. and Fujii, N., "*Convenient one-pot synthesis of cystine-containing peptides using the trimethylsilyl chloride-dimethyl sulfoxide/trifluoroacetic acid system and its application to the synthesis of bifunctional anti-HIV compounds.*" J. Chem. Soc., Perkin Trans., **1998**, *1*, 495.

157. Matsumoto, H.; Hamawaki, T.; Ota, H.; Kimura, T.; Goto, T.; Sano, K.; Hayashi, Y. and Kiso, Y., "*'Double-drugs' - A new class of prodrug form of an HIV protease inhibitor conjugated with a reverse transcriptase inhibitor by a spontaneously cleavable linker.*" Bioorg. Med. Chem. Lett., **2000**, *10*, 1227-1231.

158. Matsumoto, H.; Sohma, Y.; Kimura, T.; Hayashi, Y. and Kiso, Y., "*Controlled drug release: New water-soluble prodrugs of an HIV protease inhibitor.*" Bioorg. Med. Chem. Lett., **2001**, *11*, 605-609.

159. Hill, J. W., "*Studies on polymrization and ring formation. VI. adipic anhydride.*" J. Am. Chem. Soc., **1930**, *52*, 4110-4114.

160. Engvall, E. and Perlmann, P., "*Enzyme-linked immunosorbent assay (ELISA). Quantitative assay of immunoglobulin G.*" Immunochem., **1971**, *8*(9), 871-874.

161. Crowther, J. R., "*ELISA: Theory and Practice*", in *Methods in Molecular Biology.* **1995**, Humana Press: Totowa, New Jersey.

162. Weislow, O. S.; Kiser, R.; Fine, D. L.; Bader, J.; Shoemaker, R. H. and Boyd, M. R., "*New soluble-formazan assay for HIV-1 cytopathic effects: application to high-flux screening of synthetic and natural products for AIDS-antiviral activity.*" J. Natl. Cancer Inst., **1989**, *81*(8), 577-586.

163. Bradford, M. M., "*A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding.*" Anal. Biochem., **1976**, *72*(1-2), 248-54.

164. Gulakowski, R. J.; McMahon, J. B.; Staley, P. G.; Moran, R. A. and Boyd, M. R., "*A semiautomated multiparameter approach for anti-HIV drug screening.*" J. Virol. Methods, **1991**, *33*, 87-100.